

system using a cDNA library as prey in the presence of the PPAR ligand. Also disclosed are cells that express the PPAR protein, a method for screening for drugs for insulin resistance, and preparing a medical composition for improving insulin resistance. The method of the invention is useful for screening for drugs that improve insulin resistance. The current sequence represents a human promoter sequence.

Sequence 1870 BP; 420 A; 530 C; 459 G; 461 T; 0 U; 0 Other;

100.0%; Score 1870; DB 10; Length 1870;
Local Similarity 100.0%; Pred. No. 0;
Matches 1870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 ATCTGTCTTCTAGAAAGTACCAAGCAATCTACAGAGGCTGAAATTAATCTGTTCT 60
1 ATCTGTCTTCTAGAAAGTACCAAGCAATCTACAGAGGCTGAAATTAATCTGTTCT 60
61 AGTTCTAGATTTTCACTCATGTGCTGTGAGTATGCTCTTATTTGCCAGCTCCACTG 120
61 AGTTCTAGATTTTCACTCATGTGCTGTGAGTATGCTCTTATTTGCCAGCTCCACTG 120
121 GCAAAAGTTGGCTCTCTCTAGAGCTCTTGGATGAACTGATTTCCATGCTCAATGGCCA 180
121 GCAAAAGTTGGCTCTCTCTAGAGCTCTTGGATGAACTGATTTCCATGCTCAATGGCCA 180
181 GGAATATGACTAGAGAAAGTTACATCCAGAGAGAGGAGGAGGCTGTAAATCTGAGCAG 240
181 GGAATATGACTAGAGAAAGTTACATCCAGAGAGAGGAGGAGGCTGTAAATCTGAGCAG 240
241 CATTCCTGCTATTTCTTCTGAGCAAGAGGAAATTAATCTAGAGCAATTCATATGCAAG 300
241 CATTCCTGCTATTTCTTCTGAGCAAGAGGAAATTAATCTAGAGCAATTCATATGCAAG 300
301 AAAAATCTTAGAGTCAAGAGTACTTGGAGAGAGTACAGAAATGAAAGAGAAATTAAT 360
301 AAAAATCTTAGAGTCAAGAGTACTTGGAGAGAGTACAGAAATGAAAGAGAAATTAAT 360
361 CTGGAGAAATTAATCAAGAGAGAGAAATTAATGAGAGGCTGAGGCTTTCTTCACTT 420
361 CTGGAGAAATTAATCAAGAGAGAGAAATTAATGAGAGGCTGAGGCTTTCTTCACTT 420
421 GGGTCTCAATCAGGCTCTTGAAGAGCTGATTTCTGCTGCTGCTGAGGAGGAGGCTG 480
421 GGGTCTCAATCAGGCTCTTGAAGAGCTGATTTCTGCTGCTGCTGAGGAGGAGGCTG 480
481 CAGAGATTAAGAGTCTGCTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 540
481 CAGAGATTAAGAGTCTGCTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 540
541 TTGCTCTGTGCGCCAGGCTGAGAGTCAATGATCTGCTGCTGCTGCTGCTGCTGCTG 600
541 TTGCTCTGTGCGCCAGGCTGAGAGTCAATGATCTGCTGCTGCTGCTGCTGCTGCTG 600
601 TTCCGGATTTCAAGCGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
601 TTCCGGATTTCAAGCGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
661 GCCACACACCCAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 720
661 GCCACACACCCAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 720
721 AGTACAGTGGAGATCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
721 AGTACAGTGGAGATCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
781 CTGCTCAAGTCTCTGAGTATGAGATTAAGAGTCAAGAGTCAAGTCAAGTCAAGTCA 840
781 CTGCTCAAGTCTCTGAGTATGAGATTAAGAGTCAAGAGTCAAGTCAAGTCAAGTCA 840
841 TTGTATTTTATAGAGATGCGCTTTTGGCATGTGGCATGTGGCATGTGGCATGTGGCA 900
841 TTGTATTTTATAGAGATGCGCTTTTGGCATGTGGCATGTGGCATGTGGCATGTGGCA 900

```

```

QY 901 CCTCAGATGATCCGCTGAGCTTTGGCTCCCAAGTCTGGAGTTGAGGCGTGAAGCATC 960
DB 901 CCTCAGATGATCCGCTGAGCTTTGGCTCCCAAGTCTGGAGTTGAGGCGTGAAGCATC 960
QY 961 GGGCAGGCGTGAAGTACTCTTTAGTCTGTGAAGAAGTGGGCTGAGAGAAATCAAG 1020
DB 961 GGGCAGGCGTGAAGTACTCTTTAGTCTGTGAAGAAGTGGGCTGAGAGAAATCAAG 1020
QY 1021 CTTTACATGCCATCTCTCCCTAGATGCCAGGCTTTCTCTGAGATGCTTTCTTAATTAC 1080
DB 1021 CTTTACATGCCATCTCTCCCTAGATGCCAGGCTTTCTCTGAGATGCTTTCTTAATTAC 1080
QY 1081 CATCTCTAATCAGAGGCGCTTTAAGAGATATCAGGATGATATTTTGGTCAAGA 1140
DB 1081 CATCTCTAATCAGAGGCGCTTTAAGAGATATCAGGATGATATTTTGGTCAAGA 1140
QY 1141 TAACTTTTCCCAAGACTCAGTAAAGGCTCTGATGAGCATCAGTAAAGGCAATTCAT 1200
DB 1141 TAACTTTTCCCAAGACTCAGTAAAGGCTCTGATGAGCATCAGTAAAGGCAATTCAT 1200
QY 1201 AGCAACAGGTTCTGCTGCTTTTACAGGAGCTCCAGTTGTTGGAGTGAAGCAAGCAG 1260
DB 1201 AGCAACAGGTTCTGCTGCTTTTACAGGAGCTCCAGTTGTTGGAGTGAAGCAAGCAG 1260
QY 1261 TGAAGGGAAGGCAAGCTTTGAGAGCTTCAGGCTGAGTTCAGGAGGCTTCGAGG 1320
DB 1261 TGAAGGGAAGGCAAGCTTTGAGAGCTTCAGGCTGAGTTCAGGAGGCTTCGAGG 1320
QY 1321 GGGGCCCAAGAGTCTGAATCTAGAGGCTCCAGGCTCCCAAGGCTTCAGGCTCAG 1380
DB 1321 GGGGCCCAAGAGTCTGAATCTAGAGGCTCCAGGCTCCCAAGGCTTCAGGCTCAG 1380
QY 1381 TCTGTTTACATGAGTCTGAGGCTCAAGAACTCAGGCTGAAATTCATCTCCGTGATTA 1440
DB 1381 TCTGTTTACATGAGTCTGAGGCTCAAGAACTCAGGCTGAAATTCATCTCCGTGATTA 1440
QY 1441 CCACCTTAACATCCCATCTTATGAGTCACTCTGCTGAGGAGCAATCTGATTTG 1500
DB 1441 CCACCTTAACATCCCATCTTATGAGTCACTCTGCTGAGGAGCAATCTGATTTG 1500
QY 1501 TGTGGGCGCTGATCTGATGAGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1560
DB 1501 TGTGGGCGCTGATCTGATGAGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1560
QY 1561 CTCAACCCATCAGGCTCAGATCTCTTAAGACAGAGGATTAAGTCTCCGTGAAATC 1620
DB 1561 CTCAACCCATCAGGCTCAGATCTCTTAAGACAGAGGATTAAGTCTCCGTGAAATC 1620
QY 1621 CCTCCCGCTCATATTCCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 1621 CCTCCCGCTCATATTCCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
QY 1681 CAAAGGCGCGGTAGAGAAAGTACATTTCCAGAGGAGCAAGAGAGGAGGCTTCACT 1740
DB 1681 CAAAGGCGCGGTAGAGAAAGTACATTTCCAGAGGAGCAAGAGAGGAGGCTTCACT 1740
QY 1741 ACAACCCGCGCGGTATGAGTGGAGAAAGGAGTGAAGGAGGCTCAAGGCTGAGG 1800
DB 1741 ACAACCCGCGCGGTATGAGTGGAGAAAGGAGTGAAGGAGGCTCAAGGCTGAGG 1800
QY 1801 GGGGAGGCGCTTAACCATGCTGATCTGAGAGGAGGCGGAGGCTGAGGCTGAGG 1860
DB 1801 GGGGAGGCGCTTAACCATGCTGATCTGAGAGGAGGCGGAGGCTGAGGCTGAGG 1860
QY 1861 TCTTCTGATG 1870
DB 1861 TCTTCTGATG 1870

```

RESULT 2
ABD32700/c
ID ABD32700 standard; DNA; 33362 BP.
XX

ABD32700;
18-NOV-2004 (first entry)
Human cancer-associated genomic DNA HD14-035.
Human; dg; cancer-associated protein; gene; cytoskeletal; cancer;
leukemia; lymphoma; CAP.
Homo sapiens.
MO2004074320-A2.
02-SEP-2004.
17-FEB-2004; 2004WO-US004730.
14-FEB-2003; 2003US-00367094.
14-MAR-2003; 2003US-0038838.
15-APR-2003; 2003US-00417375.
13-JUN-2003; 2003US-00461862.
15-SEP-2003; 2003US-00663431.
15-DEC-2003; 2003US-00737318.
(SAGR-) SAGRES DISCOVERY INC.
Morris DW, Morrie DW, Malandro MS;
WPI; 2004-652914/63.
New isolated cancer-associated polynucleotides and polypeptides useful
for diagnosing, preventing or treating cancers, especially lymphoma and
leukemia, or in screening for agents that modulate cancer.
claim 16; seqid 265; 310pp; English.
The invention relates to an isolated nucleic acid comprising at least 10
contiguous nucleotides of any of the 231 polynucleotide sequences given
in the specification, or its complement. The nucleic acids encode cancer-
associated proteins. Also included are an expression vector comprising
the isolated nucleic acid cited above, a host cell comprising the above
recombinant nucleic acid or expression vector, a microarray for detecting
a cancer-associated (CA) nucleic acid comprising at least one probe
comprising at least 10 contiguous nucleotides of any of the above-
mentioned nucleotide sequences, an isolated polypeptide (encoded within
an open reading frame of a CA sequence selected from any of the 95
polynucleotide sequences as mentioned in the specification, or its
complement), an isolated antibody, (or its antigen binding fragment) that
binds to the above polypeptide, a hybridoma that produces the above
monoclonal antibody, a pharmaceutical composition comprising the above
antibody and a pharmaceutical excipient, a kit for detecting cancer
cells (comprising the antibody cited above, methods for diagnosing cancer
or for detecting the presence or absence of cancer cells in an
individual, a method for inhibiting growth of cancer cells in an
individual, a method for delivering a therapeutic agent to cancer cells
in an individual, an electronic library comprising the above
polynucleotide or polypeptide (or their fragments), methods of screening
for anticancer activity or for a bioactive agent capable of modulating
the activity of a CA protein (CAP), methods for detecting cancer
associated with expression of a polypeptide in a test cell sample, a
method for treating cancers and a method for inhibiting the expression of
CA gene in a cell. The composition and methods are useful for detecting,
diagnosing, preventing and treating cancers, especially lymphoma and
leukemia. These may also be used in screening for agents that modulate
cancer. The present sequence is a human CAP genomic sequence. Note: The
sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences

Sequence 33362 BP; 8042 A; 8668 C; 8695 G; 7957 T; 0 U; 0 Other;
Very Match 17.7%; Score 330.8; DB 13; Length 33362;
Int Local Similarity 83.0%; Pred. No. 1.1e-74;
Matches 377; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 517 TTTCTTTTGTGAGACAGAGCTTGTCTGTGCGCCAGAGCTGAGTGCAGTGA 576
DB 12186 TCTCTTTTGTGAGACAGAGCTTGTCTGTGCGCCAGAGCTGAGTGCAGTGA 12127
QY 577 TCTCTGCGCCAGTGCAGTCTTGTCTGTGCGCCAGAGCTGAGTGCAGTGA 636
DB 12126 TCTCTGCGCCAGTGCAGTCTTGTCTGTGCGCCAGAGCTGAGTGCAGTGA 12067
QY 637 AAGTAGCTGGAGTTTACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGA 696
DB 12066 GAGTAGCTGGAGTTTACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGA 12007
QY 697 GTCTTGGCCCTGTGACCCAGAGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGA 756
DB 12006 GTCTTGGCCCTGTGACCCAGAGCTGAGTGCAGTGCAGTGCAGTGCAGTGA 11947
QY 757 ACCTCCCGGGTTTCAAGCAATTTCTGCTCAAGTCTTCTGAGTGCAGTGA 816
DB 11946 GCCGCCCGGGTTTCAAGCAATTTCTGCTCAAGTCTTCTGAGTGCAGTGA 11887
QY 817 TGACCTTCAAGTTCAGTCTAATTTTGTATTTTGTATGAGTGCAGTGCAGTGA 876
DB 11886 TGACCTTCAAGTTCAGTCTAATTTTGTATTTTGTATGAGTGCAGTGCAGTGA 11827
QY 877 GCCATGCTAGTCTGAGACCCCGAGCTTCAAGTGCAGTGCAGTGCAGTGA 936
DB 11826 GCAAGGCTGATCTTGAACCTCTGACCTTCAAGTGCAGTGCAGTGCAGTGA 11767
QY 937 CTGGGATTGCAAGCGTGAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGA 970
DB 11766 CTGGGATTGCAAGCGTGAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGA 11733
RESULT 3
ACN4954
ID ACN4954 standard; DNA; 31116 BP.
XX ACN4954;
AC ACN4954;
XX 18-NOV-2004 (first entry)
DT Human genomic sequence hCG38622.
XX Human genomic sequence hCG38622.
DB Human genomic sequence hCG38622.
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
OS Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX Homo sapiens.
XX MO2003073826-A2.
PN MO2003073826-A2.
XX 12-SEP-2003.
PD 12-SEP-2003.
XX 28-FEB-2003; 2003WO-US006235.
PF 28-FEB-2003; 2003WO-US006235.
XX 01-MAR-2002; 2002US-00087192.
PR 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
PA (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
PI Morris DW;
XX WPI; 2003-328604/31.
DR WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PS Claim 1; SEQ ID NO 1660; 0pp; English.
XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for

Claim 1; SEQ ID NO 106; 2304bp; English.

The invention relates to a novel recombinant nucleic acid comprising a nucleotide sequence selected from any of the 660 sequences fully defined in the specification. A polynucleotide of the invention has cytostatic activity, and may have a use in gene therapy, or in a vaccine. The recombinant nucleic acids and polypeptides are useful for treating carcinomas, e.g. lymphomas, cancers, neoplasms, adenocarcinoma, and sarcomas. The present sequence represents a human gene of the invention.

Sequence 58822 BP; 14199 A; 14815 C; 15625 G; 13656 T; 0 U; 467 Other;

17.4%; Score 324.8; DB 10; Length 58822;

Local Similarity 77.5%; Pred. No. 4.8e-73; Indels 7; Gaps 1;

Matches 409; Conservative 0; Mismatches 112;

502 TTTTCTTTTCTTTCTTTCTTTTGTGAGACAGAGTCTGCTGTCGCGCCAGGCTG 561
58246 TTTTCTTTTCTTTCTTTCTTTTGTGAGACAGAGTCTGCTGTCGCGCCAGGATG 58187
562 AGTGCAGTGGCAGATCTCTGCGCCAGTCACTCTGCTCTCCGATTCAGAGATTCTC 621
58186 AATGACGTGGTGTGATCTGGCTCAAGACCTCTGCTCTCCAGGTTCAAGTATTTCTC 58127
622 CTGCTCAGCTCTCCCAAGTACGTGGGATTAAGGTGACGCGCCACACACCCAGC----- 675
58126 CTGCTCAGCTCTCCCAAGTACGTGGGATTAAGGTGACGCGCGCTAATTT 58067
676 TTTTCTTTTCTTTTGTGAGACAGAGTCTGCGCTGTCACCCAGGCTGAGTACAGTGGCATG 734
58066 TTTTCTTTTCTTTTGTGAGATGAGAGTCTGCTGTCGCGCCAGGCTGAGTGGCATG 58007
735 ATCTCAGTTCACTGCGACCTCCACTCCGCGGTTCAAGCAATTTCTGCTCAGTCTCC 794
58006 ATCTCAGTTCACTGCGACCTCCACTCCAGGTTCAAGGATTTCTGCTCAGTCTCC 57947
795 TGAGTAGCTAGATTAAGAGAGTCACTCCAGTTCAGTATTTTGTATTTTAACTA 854
57946 CGAGTAGCTAGATTAAGAGAGTCACTCCAGTTCAGTATTTTGTATTTTAACTA 57887
855 GAGATGCGCTTTTGGCAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 914
57886 GAGATGCGCTTTTGGCAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 57827
915 CTGCGCTTGGCTCTCCCAAGAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
57826 CCGTCTTGGCTCTCCCAAGAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 57767
975 CTACTCTTAACTCTCTGAGAAAGTCTGCGGCTCAAGAAATCAAGCT 1022
57766 TTGCTTTCTTTTCTTTTCTGAGACAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 57719

12-FEB-2004 (first entry)

Human TCOF1 gene genomic DNA sequence.

cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy; lymphoma; breast cancer; prostate cancer; leukaemia; ds; human; TCOF1.

Homo sapiens.

WO2003039484-A2.

15-MAY-2003.

08-NOV-2002; 2002MO-US036071.

XX 08-NOV-2001; 2001US-00052482.
PR (SAGR-) SAGRES DISCOVERY.
XX Morris DW, Engelhard EK.
XX WPI; 2003-441462/41.
XX
XX New carcinoma associated nucleic acids and proteins, useful for screening
XX drug candidates, or for diagnosing and treating carcinomas, e.g.
XX Lymphoma, breast cancer, prostate cancer or leukemia.
XX
XX Claim 1; SEQ ID NO 46; 793bp; English.
PS
XX
XX This invention relates to novel recombinant nucleic acids for use in
XX diagnosis and treatment of cancer, especially carcinomas, as well as the
XX use of compositions in screening methods. The compositions of the
XX invention may have cytostatic activity whilst the disclosed sequences may
XX be useful for gene therapy. The carcinoma associated nucleic acids and
XX proteins are useful for diagnosing and treating carcinomas, for example
XX lymphoma, breast cancer, prostate cancer or leukaemia, or for screening
XX drug candidates or bioactive agents capable of binding to, or modulating
XX the activity of, a carcinoma associated protein. The present sequence is
XX the genomic DNA sequence of the human TCOF1 gene which is a carcinoma
XX associated gene of the invention.
XX
XX Sequence 58822 BP; 14199 A; 14881 C; 15619 G; 13656 T; 0 U; 467 Other;
XX
XX Query Match 17.4%; Score 324.8; DB 10; Length 58822;
XX Best Local Similarity 77.5%; Pred. No. 4.8e-73;
XX Matches 409; Conservative 0; Mismatches 112; Indels 7; Gaps 1;
XX
XX 502 TTTTCTTTTCTTTCTTTCTTTTGTGAGACAGAGTCTGCTGTCGCGCCAGGCTG 561
XX 58246 TTTTCTTTTCTTTCTTTCTTTTGTGAGACAGAGTCTGCTGTCGCGCCAGGATG 58187
XX 562 AGTGCAGTGGCAGATCTCTGCGCCAGTCACTCTGCTCTCCGATTCAGAGATTCTC 621
XX 58186 AATGACGTGGTGTGATCTGGCTCAAGACCTCTGCTCTCCAGGTTCAAGTATTTCTC 58127
XX 622 CTGCTCAGCTCTCCCAAGTACGTGGGATTAAGGTGACGCGCGCTAATTT 58067
XX 58126 CTGCTCAGCTCTCCCAAGTACGTGGGATTAAGGTGACGCGCGCTAATTT 58007
XX 676 TTTTCTTTTCTTTTGTGAGACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 734
XX 58066 TTTTCTTTTCTTTTGTGAGATGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 58007
XX 735 ATCTCAGTTCACTGCGACCTCCACTCCGCGGTTCAAGCAATTTCTGCTCAGTCTCC 794
XX 58006 ATCTCAGTTCACTGCGACCTCCACTCCAGGTTCAAGGATTTCTGCTCAGTCTCC 57947
XX 795 TGAGTAGCTAGATTAAGAGAGTCACTCCAGTTCAGTATTTTGTATTTTAACTA 854
XX 57946 CGAGTAGCTAGATTAAGAGAGTCACTCCAGTTCAGTATTTTGTATTTTAACTA 57887
XX 855 GAGATGCGCTTTTGGCAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 914
XX 57886 GAGATGCGCTTTTGGCAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 57827
XX 915 CTGCGCTTGGCTCTCCCAAGAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
XX 57826 CCGTCTTGGCTCTCCCAAGAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 57767
XX 975 CTACTCTTAACTCTCTGAGAAAGTCTGCGGCTCAAGAAATCAAGCT 1022
XX 57766 TTGCTTTCTTTTCTTTTCTGAGACAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 57719

RESULT 8
AAK78433/C
ID AAK78433 standard; DNA; 24934 BP.

AAK78433;
07-NOV-2001 (first entry)
Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33245.
Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
cytostatic; gene therapy; vaccine; metastasis; ds.
Homo sapiens.
MO200157182-A2.
09-AUG-2001.
17-JAN-2001; 2001MO-US001354.
31-JAN-2000; 2000US-0179065P.
04-FEB-2000; 2000US-0180628P.
24-FEB-2000; 2000US-0184664P.
02-MAR-2000; 2000US-0186350P.
16-MAR-2000; 2000US-0189874P.
17-MAR-2000; 2000US-0190076P.
18-APR-2000; 2000US-0198123P.
19-MAY-2000; 2000US-0205515P.
07-JUN-2000; 2000US-0209467P.
28-JUN-2000; 2000US-0214886P.
30-JUN-2000; 2000US-0215135P.
07-JUL-2000; 2000US-0216647P.
07-JUL-2000; 2000US-0216880P.
11-JUL-2000; 2000US-0217487P.
11-JUL-2000; 2000US-0217496P.
14-JUL-2000; 2000US-0218290P.
26-JUL-2000; 2000US-0220963P.
26-JUL-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225213P.
14-AUG-2000; 2000US-0225214P.
14-AUG-2000; 2000US-0225266P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225268P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
14-AUG-2000; 2000US-0225759P.
16-AUG-2000; 2000US-0226279P.
22-AUG-2000; 2000US-0226681P.
22-AUG-2000; 2000US-0226686P.
22-AUG-2000; 2000US-0227182P.
23-AUG-2000; 2000US-0227009P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
06-SEP-2000; 2000US-0230437P.
06-SEP-2000; 2000US-0230438P.
08-SEP-2000; 2000US-0231242P.
08-SEP-2000; 2000US-0231243P.
08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232387P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240360P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250310P.
PR 01-DEC-2000; 2000US-0250311P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.

16-MAR-2000; 2000US-0189874P.
 17-MAR-2000; 2000US-0190076P.
 18-APR-2000; 2000US-0198133P.
 19-MAY-2000; 2000US-0205515P.
 07-JUN-2000; 2000US-0209467P.
 28-JUN-2000; 2000US-0214886P.
 30-JUN-2000; 2000US-0215135P.
 07-JUL-2000; 2000US-0216647P.
 07-JUL-2000; 2000US-0216880P.
 11-JUL-2000; 2000US-0217487P.
 11-JUL-2000; 2000US-0217496P.
 14-JUL-2000; 2000US-0218290P.
 26-JUL-2000; 2000US-0220963P.
 26-JUL-2000; 2000US-0220964P.
 14-AUG-2000; 2000US-0224518P.
 14-AUG-2000; 2000US-0224519P.
 14-AUG-2000; 2000US-0225213P.
 14-AUG-2000; 2000US-0225214P.
 14-AUG-2000; 2000US-0225266P.
 14-AUG-2000; 2000US-0225267P.
 14-AUG-2000; 2000US-0225268P.
 14-AUG-2000; 2000US-0225270P.
 14-AUG-2000; 2000US-0225447P.
 14-AUG-2000; 2000US-0225757P.
 14-AUG-2000; 2000US-0225758P.
 14-AUG-2000; 2000US-0225759P.
 18-AUG-2000; 2000US-0226279P.
 22-AUG-2000; 2000US-0226681P.
 22-AUG-2000; 2000US-0226686P.
 22-AUG-2000; 2000US-0227182P.
 23-AUG-2000; 2000US-0227009P.
 30-AUG-2000; 2000US-0228924P.
 01-SEP-2000; 2000US-0229287P.
 01-SEP-2000; 2000US-0229343P.
 01-SEP-2000; 2000US-0229344P.
 01-SEP-2000; 2000US-0229345P.
 05-SEP-2000; 2000US-0229509P.
 05-SEP-2000; 2000US-0229513P.
 06-SEP-2000; 2000US-0230437P.
 06-SEP-2000; 2000US-0230438P.
 08-SEP-2000; 2000US-0231242P.
 08-SEP-2000; 2000US-0231243P.
 08-SEP-2000; 2000US-0231244P.
 08-SEP-2000; 2000US-0231413P.
 08-SEP-2000; 2000US-0231414P.
 08-SEP-2000; 2000US-0232080P.
 08-SEP-2000; 2000US-0232081P.
 12-SEP-2000; 2000US-0231968P.
 14-SEP-2000; 2000US-0232397P.
 14-SEP-2000; 2000US-0232398P.
 14-SEP-2000; 2000US-0232399P.
 14-SEP-2000; 2000US-0232400P.
 14-SEP-2000; 2000US-0232401P.
 14-SEP-2000; 2000US-0233063P.
 14-SEP-2000; 2000US-0233064P.
 21-SEP-2000; 2000US-0234223P.
 21-SEP-2000; 2000US-0234274P.
 25-SEP-2000; 2000US-0234997P.
 25-SEP-2000; 2000US-0234998P.
 26-SEP-2000; 2000US-0235484P.
 27-SEP-2000; 2000US-0235834P.
 27-SEP-2000; 2000US-0235836P.
 29-SEP-2000; 2000US-0236337P.
 29-SEP-2000; 2000US-0236357P.
 29-SEP-2000; 2000US-0236368P.
 29-SEP-2000; 2000US-0236369P.
 29-SEP-2000; 2000US-0236370P.
 02-OCT-2000; 2000US-0236802P.
 02-OCT-2000; 2000US-0237037P.
 02-OCT-2000; 2000US-0237038P.
 02-OCT-2000; 2000US-0237039P.
 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 20-OCT-2000; 2000US-0242221P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246529P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 01-DEC-2000; 2000US-0251160P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251899P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-541565/60.
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system cancers
 XX and metastases.
 PS Disclosure; SEQ ID NO 11292; 1701pp + Sequence Listing; English.
 XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins

(ABR14678-ABR18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 10663 BP; 2921 A; 2087 C; 2099 G; 3555 T; 0 U; 1 Other;

Query Match 17.3%; Score 323.4; DB 5; Length 10663;
Local Similarity 81.3%; Pred. No. 5.6e-73;
Matches 391; Conservative 0; Mismatches 81; Indels 9; Gaps 1;

```

502 TTTTCTTTTCTTTCTTTCTTTTGTGAGACAGTCTGTCTGTGCGCCAGGCTGG 561
5384 TTTTCTTTTCTTTCTTTCTTTTGTGAGACAGTCTGTCTGTGCGCCAGGCTGG 5443
562 AGTGCAGTGGCATGATCTGCGCCAGTGCACCTCTGCGGATTAAGGATCTTC 621
5444 AGTGTGTGTGTGATCTGTGCGGCTGCACTCTGCTCTCAAGTTCAAGTATCTC 5503
622 CTGCTCAGCCTCCCAAGTAGCTGGATTACAGTGCACGCCACACCCAGC----- 675
5504 CTGCTCAGCCTCTGAGTAGTAGTACTACAGGACAGTGCACGCCACGCTAATT 5563
676 ---TTTCTTTTATTTGAGACAGAGTCTGCTGTCACCCAGGCTGGAGTACAGTGA 732
5564 TTGTTTTTGTGTTTGTGAGACAGAGTCTCACTGTGCGGACAGGCTGGAGTGA 5623
733 TGATCTCAGTTCACCTGAGCCTCCACCTCCCGGGTTCAGCAATTTCTCTGCTCAGTCT 792
5624 CATCTCAGTCTCAGTCACTGCACTCTGCGGCTCCGCGGCTCAACATTTCTCTGCTCAGTCT 5683
793 CCTGAGTAGTAGTATTAAGAGTGCACCTTCACGTTCAAGTAAATTTTGTATTTTAA 852
5684 CCCAAGTAGTGGATTACAGGTGTGTGACCATGCGCGCTAATTTTGTATTTTAA 5743
853 TAGAGTAGGCTTTTGGCAGTGTGGCAGTCTGAGAACCCCGGACCTCAGGTATC 912
5744 TAGAGTAGGCTTTTGGCAGTGTGGCAGTCTGAGAACCTTGAACCTTGAACCAAGTAAATC 5803
913 CGTGGCCTTTGAGCTCCCAAGGTGTGGAGTTCAGGCGTGAAGCCATGCGCCAGGCTGG 972
5804 CGTGGCCTTTGAGCTCCCAAGGTGTGGAGTTCAGGCGTGAAGCCATGCGCTTCCGCGCTTA 5863
973 A 973
5864 A 5864

```

ACN44538; standard; DNA; 296405 BP.

18-NOV-2004 (first entry)

Human genomic sequence hCG16326.

Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

Homo sapiens.

XX WO2003073826-A2.
PN 12-SEP-2003.
XX 28-FEB-2003; 2003MO-US006235.
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX Claim 1; SEQ ID NO 1036; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published

Sequence 296405 BP; 82162 A; 54238 C; 57164 G; 90624 T; 0 U; 12217 Other;

Query Match 17.3%; Score 322.6; DB 11; Length 296405;
Best Local Similarity 79.7%; Pred. No. 3.5e-72;
Matches 396; Conservative 0; Mismatches 94; Indels 7; Gaps 1;

```

488 TGAGCTGCTGCTGCTTTTCTTTCTTTTGTGAGACAGAGTCTGCTCT 547
117715 TGTGATGATGCAATATGAAATCTTTTGTGTTTATTTTGGATGAGTCACTCT 117774
548 GTGCGCCAGGCTGAGTGCAGTGCATGATCTCTGCCATGCACTCTGCTCCCGGA 607
117775 GTTACCCAGGCTGAGTGCAGTGCATGATCTCACTGCTGCTCAACTCTGCTCCCGG 117834
608 TTCAAGCATTTCTCTGCTCTGAGCTCCCAATAGTGGAGTTACAGTGCACAGCAGCA 667
117835 TTCCAGCATTTCTCTGCTCTGAGCTCCCAATAGTGGAGTTACAGGCACTGCGCACA 117894
668 CACCCAGCT-----TTTCTTTTATTTGAGACAGAGTCTTCCCTGTGACCCAGGCTGG 720
117895 TGTCTAGCTAATTTTGTGTTTGTGTTTGTGAGACAGAGTCTTACTGTGTGCGCAGGCTGG 117954
721 AGTACAGTGGCATGATCTGATCTGACCTGACCTTCACCTCCCGGGTTCAAGCAATCTC 780
117955 AGTGCAGTGGCAATATCTGCGCTGCTGCAACTCCACCTCCCGGGTTCAAGCAATCTCT 118014
781 CTGCTCAGTCTCTGAGTGCAGTGCATGATCTGAGTGCACCTCCACCTGTCAGTAAATTT 840
118015 CTGCTCAGTCTCTGAGTGCAGTGCATGATCTGAGTGCACCTCCACCTGTCAGTAAATTT 118074
841 TTGTATTTTATGAGATGAGTGGCTTTGCGCATGTTGGCCATGATCTGTAACCCCGGA 900
118075 TTGTATTTTATGAGATGAGTGGCTTTGCGCATGTTGGCCATGATCTGTAACCCCGGA 118134
901 CCTCAGTGAATCGCTGAGCTTTGAGCTCCCAAGGTGTGGAGTTCAGGCGTGAAGCCATC 960
118135 CCTCAGTGAATCGCTGAGCTTTGAGCTCCCAAGGTGTGGAGTTCAGGCGTGAAGCCATC 118194

```

961 GCGCCAGCCTGAGCTA 977
 118195 CTGCGCGCCTGAATA 118211

17-13
 7799/C
 AAS32799 standard; DNA; 4957 BP.

AAS32799;

17-DEC-2001 (first entry)

Human genomic DNA for novel endocrine antigen, SEQ ID NO 753.

Human; endocrine antigen; ds; cytosolic; antifertility; antidiabetic;
 thyroid-active; adrenal-active; androgenic; gastric; gene therapy;
 antitense-therapy; antibody; endocrine disorder; hormone imbalance;
 reproductive disorder; endocrine cancer; pancreatic disorder;
 diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;
 hyperthyroidism; hypothalamic disorder; vanishing testes syndrome.

Homo sapiens.

MO200155319-A2.

02-AUG-2001.

17-JAN-2001; 2001MO-US001335.

31-JAN-2000; 2000US-0179065P.
 04-FEB-2000; 2000US-0180628P.
 24-FEB-2000; 2000US-0184664P.
 02-MAR-2000; 2000US-0186350P.
 16-MAR-2000; 2000US-0189874P.
 17-MAR-2000; 2000US-0190076P.
 18-APR-2000; 2000US-0198133P.
 19-MAY-2000; 2000US-0205515P.
 07-JUN-2000; 2000US-0209467P.
 28-JUN-2000; 2000US-0214886P.
 30-JUN-2000; 2000US-0215135P.
 07-JUL-2000; 2000US-0216647P.
 07-JUL-2000; 2000US-0216880P.
 11-JUL-2000; 2000US-0217487P.
 11-JUL-2000; 2000US-0217496P.
 14-JUL-2000; 2000US-0218290P.
 26-JUL-2000; 2000US-0220963P.
 26-JUL-2000; 2000US-0220964P.
 14-AUG-2000; 2000US-0224518P.
 14-AUG-2000; 2000US-0224519P.
 14-AUG-2000; 2000US-0225213P.
 14-AUG-2000; 2000US-0225214P.
 14-AUG-2000; 2000US-0225266P.
 14-AUG-2000; 2000US-0225267P.
 14-AUG-2000; 2000US-0225268P.
 14-AUG-2000; 2000US-0225447P.
 14-AUG-2000; 2000US-0225757P.
 14-AUG-2000; 2000US-0225758P.
 14-AUG-2000; 2000US-0225759P.
 18-AUG-2000; 2000US-0226279P.
 22-AUG-2000; 2000US-0226681P.
 22-AUG-2000; 2000US-0226686P.
 23-AUG-2000; 2000US-0227182P.
 30-AUG-2000; 2000US-0227009P.
 01-SEP-2000; 2000US-0228924P.
 01-SEP-2000; 2000US-0229287P.
 01-SEP-2000; 2000US-0229343P.
 01-SEP-2000; 2000US-0229344P.
 01-SEP-2000; 2000US-0229345P.
 05-SEP-2000; 2000US-0229509P.
 05-SEP-2000; 2000US-0229513P.
 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241211P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246533P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.

14-AUG-2000	2000US-0225758P	PR	08-NOV-2000	2000US-0246610P	PR
14-AUG-2000	2000US-0225759P	PR	08-NOV-2000	2000US-0246611P	PR
18-AUG-2000	2000US-0226279P	PR	08-NOV-2000	2000US-0246613P	PR
22-AUG-2000	2000US-0226681P	PR	17-NOV-2000	2000US-0249207P	PR
22-AUG-2000	2000US-0226688P	PR	17-NOV-2000	2000US-0249208P	PR
22-AUG-2000	2000US-0227182P	PR	17-NOV-2000	2000US-0249209P	PR
23-AUG-2000	2000US-0227009P	PR	17-NOV-2000	2000US-0249210P	PR
30-AUG-2000	2000US-0228924P	PR	17-NOV-2000	2000US-0249211P	PR
01-SEP-2000	2000US-0229287P	PR	17-NOV-2000	2000US-0249212P	PR
01-SEP-2000	2000US-0229343P	PR	17-NOV-2000	2000US-0249213P	PR
01-SEP-2000	2000US-0229344P	PR	17-NOV-2000	2000US-0249214P	PR
01-SEP-2000	2000US-0229345P	PR	17-NOV-2000	2000US-0249215P	PR
05-SEP-2000	2000US-0229509P	PR	17-NOV-2000	2000US-0249216P	PR
05-SEP-2000	2000US-0229513P	PR	17-NOV-2000	2000US-0249217P	PR
06-SEP-2000	2000US-0230437P	PR	17-NOV-2000	2000US-0249218P	PR
06-SEP-2000	2000US-0230438P	PR	17-NOV-2000	2000US-0249244P	PR
08-SEP-2000	2000US-0231242P	PR	17-NOV-2000	2000US-0249245P	PR
08-SEP-2000	2000US-0231243P	PR	17-NOV-2000	2000US-0249264P	PR
08-SEP-2000	2000US-0231244P	PR	17-NOV-2000	2000US-0249265P	PR
08-SEP-2000	2000US-0231413P	PR	17-NOV-2000	2000US-0249297P	PR
08-SEP-2000	2000US-0231414P	PR	17-NOV-2000	2000US-0249299P	PR
08-SEP-2000	2000US-0232080P	PR	17-NOV-2000	2000US-0249300P	PR
12-SEP-2000	2000US-0232081P	PR	01-DEC-2000	2000US-0250160P	PR
14-SEP-2000	2000US-0232397P	PR	01-DEC-2000	2000US-0250319P	PR
14-SEP-2000	2000US-0232398P	PR	05-DEC-2000	2000US-0251030P	PR
14-SEP-2000	2000US-0232399P	PR	05-DEC-2000	2000US-0251988P	PR
14-SEP-2000	2000US-0232400P	PR	05-DEC-2000	2000US-0256719P	PR
14-SEP-2000	2000US-0232401P	PR	06-DEC-2000	2000US-0251479P	PR
14-SEP-2000	2000US-0233063P	PR	08-DEC-2000	2000US-0251868P	PR
14-SEP-2000	2000US-0233064P	PR	08-DEC-2000	2000US-0251869P	PR
14-SEP-2000	2000US-0233065P	PR	08-DEC-2000	2000US-0251989P	PR
21-SEP-2000	2000US-0234223P	PR	08-DEC-2000	2000US-0251990P	PR
21-SEP-2000	2000US-0234274P	PR	11-DEC-2000	2000US-0254097P	PR
25-SEP-2000	2000US-0234997P	PR	05-JAN-2001	2001US-0259678P	PR
25-SEP-2000	2000US-0234998P	XX			
26-SEP-2000	2000US-0234984P	PA			
27-SEP-2000	2000US-0235634P	XX			
27-SEP-2000	2000US-0235636P	XX			
29-SEP-2000	2000US-0236327P	XX			
29-SEP-2000	2000US-0236367P	DR			
29-SEP-2000	2000US-0236368P	XX			
29-SEP-2000	2000US-0236369P	XX			
29-SEP-2000	2000US-0236370P	PT			
02-OCT-2000	2000US-0236802P	PT			
02-OCT-2000	2000US-0237037P	PT			
02-OCT-2000	2000US-0237038P	XX			
02-OCT-2000	2000US-0237039P	PS			
02-OCT-2000	2000US-0237040P	XX			
13-OCT-2000	2000US-0239935P	CC			
13-OCT-2000	2000US-0239937P	CC			
20-OCT-2000	2000US-0240960P	CC			
20-OCT-2000	2000US-0241212P	CC			
20-OCT-2000	2000US-0241785P	CC			
20-OCT-2000	2000US-0241786P	CC			
20-OCT-2000	2000US-0241787P	CC			
20-OCT-2000</					

08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249219P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251858P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-45726/49.
Isolated polypeptide for treating, preventing and prognosing disorders related to the endocrine system including endocrine disorders, reproductive disorder, and gastrointestinal disorders and also for testing and detection e.g. diagnosis.
Disclosure; SEQ ID NO 752; 558bp; English.
The invention relates to cDNAs encoding novel human endocrine antigens or a fragment having biological activity, a domain, an epitope, full length protein, variant, allelic variant or a species homologue of the cDNA/antigen. The DNAs and polypeptides are useful for preventing, treating or ameliorating a medical condition when administered (e.g. by gene therapy or antisense-therapy). Identifying mutations in the genes coding for the antigens is useful for diagnosing a pathological condition or a susceptibility to a pathological condition. The DNAs, antigens and antibodies raised against the antigens useful for treating, preventing and/or prognosing disorders related to the endocrine system or hormone imbalance or reproductive disorders, cancers of endocrine tissues, disorders of the pancreas (e.g. diabetes mellitus), the adrenal glands (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the hypothalamus and testes (e.g. vanishing testes syndrome), many examples of diseases and disorders are given in the specification. The present sequence is genomic DNA fragment form a gene encoding an endocrine antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences
Sequence 4961 BP; 1245 A; 1244 C; 1408 G; 1064 T; 0 U; 0 Other;
Query Match 17.1%; Score 320.4; DB 4; Length 4961;
Best Local Similarity 82.8%; Pred. No. 2,5e-72;

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

nucleic - nucleic search, using sw model

on: June 21, 2006, 21:27:40 ; Search time 10368 Seconds
(without alignments)
11533.727 Million cell updates/sec

US-10-502-279-26

ect score: 1870
ence: 1 atctgtctcctagaagctac.....catcccgatcctctag 1870

ing table: IDENTITY NUC
Gapop 10'0 , Gapext 1.0

ched: 6366136 sege, 31973710525 residues

1 number of hits satisfying chosen parameters: 12732272

num DB seq length: 0

num DB seq length: 2000000000

-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ibase : GenEmb1.*

1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_to:*
7: gb_scs:*
8: gb_ey:*
9: gb_un:*
10: gb_vi:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_cm:*
15: gb_da:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

to.	Score	Query Match	Length	ID	Description
1	1854.2	99.2	217470	5	AC040162 Homo sapi
2	338.8	18.1	61954	12	AC120915 Homo sapi
3	338.8	18.1	63187	5	AL355477 Human DNA
4	338.8	18.1	115710	12	AL139015 Homo sapi
5	338.8	18.1	149138	12	AC026936 Homo sapi
6	335	17.9	128829	5	AC127002 Homo sapi
7	333.4	17.8	201508	12	AC026290 Homo sapi
8	333	17.8	191353	12	AC090320 Homo sapi
9	332.6	17.8	184349	5	AC113189 Homo sapi
10	332.6	17.8	187710	5	CNS01DVI Human chr
11	332	17.8	217260	5	AC145375 Pan trogl
12	331.2	17.7	88176	12	AL350202_09 Continuation (10 o
13	331.2	17.7	185353	5	AL162732 Human DNA
14	331	17.7	101029	5	AC026954 Homo sapi
15	331	17.7	121017	5	AC087388 Homo sapi
16	331	17.7	166941	12	AC008049 Homo sapi
17	331	17.7	222871	12	AC108094 Homo sapi
18	330.8	17.7	33362	2	C0869844 Sequence

C	19	330.8	17.7	92797	5	AF038458 Homo sapi
C	20	330.4	17.7	50303	5	DQ374394 Homo sapi
C	21	330.4	17.7	162991	5	AC027682 Homo sapi
C	22	330.4	17.7	179641	5	AC098591 Homo sapi
C	23	330.4	17.7	200618	5	AC009061 Homo sapi
C	24	329.8	17.6	39978	5	AC005197 Homo sapi
C	25	329.8	17.6	14635	5	AC090239 Homo sapi
C	26	329.8	17.6	186044	12	AC087674 Homo sapi
C	27	329.4	17.6	118684	5	AC025335 Homo sapi
C	28	329	17.6	167996	12	AC164921 Pan trogl
C	29	328.8	17.6	188466	12	AP001796 Homo sapi
C	30	328.6	17.6	91672	5	AC006468 Homo sapi
C	31	328.6	17.6	218465	12	AC127470 Pan trogl
C	32	328.4	17.6	18868	5	AC023232 Homo sapi
C	33	328	17.5	57013	12	AC015769 Homo sapi
C	34	327.6	17.5	125756	5	AL590369 Human DNA
C	35	327.6	17.5	185511	12	AC007779 Homo sapi
C	36	327.4	17.5	16372	5	AC103828 Homo sapi
C	37	327.4	17.5	145450	12	AC068352 Homo sapi
C	38	327.4	17.5	153936	5	AP001207 Homo sapi
C	39	327.4	17.5	166471	12	AC091052 Homo sapi
C	40	327	17.5	163542	12	AC129071 Homo sapi
C	41	326.4	17.5	125445	5	AC078795 Homo sapi
C	42	326.4	17.5	131682	5	AL672277 Human DNA
C	43	326.4	17.5	148152	12	AC022200 Homo sapi
C	44	326.4	17.5	154563	5	CR936360 Human DNA
C	45	326.2	17.4	184857	12	AP002503 Homo sapi

ALIGNMENTS

RESULT 1	AC040162	217470 bp	DNA	linear	PRI 28-MAR-2003
LOCUS	Homo sapiens chromosome 16 clone CTC-479C5, complete sequence.				
DEFINITION	AC040162				
ACCESSION	AC040162.5	GI:29336195			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.				
AUTHORS	1 (bases 1 to 217470)				
TITLE	Alamos National Laboratory.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (11-APR-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
AUTHORS	3 (bases 1 to 217470)				
TITLE	DOE Joint Genome Institute.				
JOURNAL	Submitted (03-APR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
REFERENCE	4 (bases 1 to 217470)				
AUTHORS	Alamos National Laboratory.				
TITLE	DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.				
JOURNAL	Submitted (28-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
COMMENT	On Mar 28, 2003 this sequence version replaced gi:19909394. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory www.sngc.stanford.edu Quality: Phrap Quality >=40 100% of Sequence; Estimated Total Number of Errors is 0.				

Location/Qualifiers
1. 217470
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTC-479C5"

ery Match 99.2%; Score 1854.2; DB 5; Length 217470;
st Local Similarity 99.8%; Pred. No. 0;
ches 1867; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

1 ATCTGTCTTGAAGAAAGTACCGACGACAACTTACAGGGTCTGAAATTAAGTCTTCT 60
16606 ATCTGTCTTGAAGAAAGTACCGACGACAACTTACAGGGTCTGAAATTAAGTCTTCT 16665
61 AGTTCTAGATTTTCACTCTTGTCTTGTGGTAAAGTCCCTTATTTGCGAGCTCCACTG 120
16666 AGTTCTAGATTTTCACTCTTGTCTTGTGGTAAAGTCCCTTATTTGCGAGCTCCACTG 166725
121 GCAAGTTGGCTCTCTCTGAGCTCTTGGATGAACTGATTTTCAATGCTCATGAGGCA 180
166726 GCAAGTTGGCTCTCTCTGAGCTCTTGGATGAACTGATTTTCAATGCTCATGAGGCA 166785
181 GGAATATGACTAGAAAGTTTACATCCAGAGAGAGGAGGCTTGGTAAATCTGAGCAG 240
166786 GGAATATGACTAGAAAGTTTACATCCAGAGAGAGGAGGCTTGGTAAATCTGAGCAG 166845
241 CATCTGCTGCTATTTCTTCTGACAGAGAGAAATTAATAGAGCAATTCATATGCAAG 300
166846 CATCTGCTGCTATTTCTTCTGACAGAGAGAAATTAATAGAGCAATTCATATGCAAG 166905
301 AAAAATTCTAGGCTCAAGAGTACTTGGAGAGAGTACAGAGAAATGAAAGAGAAATTAAC 360
166906 AAAAATTCTAGGCTCAAGAGTACTTGGAGAGAGTACAGAGAAATGAAAGAGAAATTAAC 166965
361 CTGGGAAATATTTCAAG 420
166966 CTGGGAAATATTTCAAG 167025
421 GGGTCTCAATCAGCCCTTCTAAGAGAGTATTTCTGCTGCTCTGAGGAGAGAGAGAG 480
167026 GGGTCTCAATCAGCCCTTCTAAGAGAGTATTTCTGCTGCTCTGAGGAGAGAGAGAG 167085
481 CAGAGATTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
167086 CAGAGATTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 167145
540 CTTCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
167146 CTTCGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 167205
600 CTTCGCGATTCAAGGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
167206 CTTCGCGATTCAAGGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 167265
660 GCGCAGCAGACCCAGCTTTTATTTTGAAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCT 719
167266 GCGCAGCAGACCCAGCTTTTATTTTGAAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCT 167325
720 GAGTACAGTGGCATGATCTCAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
167326 GAGTACAGTGGCATGATCTCAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 167385
780 CCTGCTCAGTCTCTGAGTAGAGATTAAGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 839
167386 CCTGCTCAGTCTCTGAGTAGAGATTAAGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 167445
840 TTTGTAATTTTGAAGAGAGTGGCTTTTGCATGTTGGCATGCTGTAAGTCTGAAACCCCG 899
167446 TTTGTAATTTTGAAGAGAGTGGCTTTTGCATGTTGGCATGCTGTAAGTCTGTAACCCCG 167505

QY 900 ACTTACAGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
DB 167506 ACTTACAGTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 167565
QY 960 GCGCAGGAGCTGAGCTACTCTGTTATGCTCTGAGAGAGCTGCGGCTCAGAGAAATCAAC 1019
DB 167566 GCGCAGGAGCTGAGCTACTCTGTTATGCTCTGAGAGAGCTGCGGCTCAGAGAAATCAAC 167625
QY 1020 GCTTACATGCCATCTCTCCCTAGTCCCAAGAGTCTTCTGAGTCTGTTATTTA 1079
DB 167626 GCTTACATGCCATCTCTCCCTAGTCCCAAGAGTCTTCTGAGTCTGTTATTTA 167685
QY 1080 CCATCTCTCAATCAGAGAGGCTTTTACAAAGATATCAGATATGATATTTTGTCAAG 1139
DB 167686 CCATCTCTCAATCAGAGAGGCTTTTACAAAGATATCAGATATGATATTTTGTCAAG 167745
QY 1140 ATATCTCTTCCCGAGAGTCTGATGAGGCTCTGATGAGCATGATAGGAGCAATTTCA 1199
DB 167746 ATATCTCTTCCCGAGAGTCTGATGAGGCTCTGATGAGCATGATAGGAGCAATTTCA 167805
QY 1200 AAGCAACAGGTTCTGCTGCTGTTTACAGGAGCTCAGTTGTTGGAAATGAGCAAGCA 1259
DB 167806 AAGCAACAGGTTCTGCTGCTGTTTACAGGAGCTCAGTTGTTGGAAATGAGCAAGCA 167865
QY 1260 GTAGAGGAGAGGAGAGCTTTCTGAGCTTCCAGGCTGATTTCTGCAAGGCTTGGG 1319
DB 167866 GTAGAGGAGAGGAGAGCTTTCTGAGCTTCCAGGCTGATTTCTGCAAGGCTTGGG 167925
QY 1320 AAGGAGGAGAGAGAGCTGAAATCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1379
DB 167926 AAGGAGGAGAGAGAGCTGAAATCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 167985
QY 1380 TTCTGCTTACAGTACAGTCAAGAACTCAGGCTGATTTTACTTCCCTGCTGATTAAG 1439
DB 167986 TTCTGCTTACAGTACAGTCAAGAACTCAGGCTGATTTTACTTCCCTGCTGATTAAG 168045
QY 1440 TCCCACTTAAATCCATCTTATTAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1499
DB 168046 TCCCACTTAAATCCATCTTATTAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 168105
QY 1500 CTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1559
DB 168106 CTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 168165
QY 1560 CCTCAGCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1619
DB 168166 CCTCAGCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 168225
QY 1620 GCGCTCCGCTCATATTCCTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1679
DB 168226 GCGCTCCGCTCATATTCCTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 168285
QY 1680 ACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1739
DB 168286 ACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 168345
QY 1740 TACACCCGCGCGGCTATGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1799
DB 168346 TACACCCGCGCGGCTATGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 168405
QY 1800 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1859
DB 168406 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 168465
QY 1860 ATCTCTGCTAG 1870
DB 168466 ATCTCTGCTAG 168476

RESULT 2
AC120915 61954 bp DNA linear HTG 23-JUL-2002
LOCUS AC120915
DEFINITION Homo sapiens, *** SEQUENCING IN PROGRESS ***, 22 unordered pieces.

```

:SESSION AC120915 GI:21903270
:ION      HTG: HTGS PHASRL
:ORDS     Homo sapiens (human)
:CE        Homo sapiens
:GANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Homidae; Homo.
:REFERENCE 1 (bases 1 to 61954)
:THORS     Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,P.R., Allen,C.,
           Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T.,
           Barbieri,J., Benton,J., Bimaga,K., Blankenburg,K., Bonini,D.,
           Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
           Buay,C., Burck,P., Burckett,C., Burrell,K.L., Byrd,N.C.,
           Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
           Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
           Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
           Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
           Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
           Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
           Barnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
           Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
           Gabili,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
           Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
           Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
           Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
           Homes,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
           Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
           Karlsson,E., Kelly,S., Khan,U., King,L., Korah,J., Kovar,C.,
           Kratovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
           Li,J., Li,Z., Licharge,O., Lien,C., Liu,J., Liu,W., Louieged,H.,
           Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
           Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
           Massey,E., Mawliny,E., Mcleod,M.P., Meador,M., Mei,G., Mezzer,M.,
           Miner,G., Miner,Z., Mitchell,T., Mohabhat,K., Morgan,M., Morris,S.,
           Moser,M., Neal,D., Newton,J., Newson,N., Nguyen,A., Nguyen,N.,
           Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogund,M., Okunnu,G.,
           Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
           Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
           Rivers,M., Rojas,A., Rojupokan,I., Rolfe,M., Ruiz,S., Savary,G.,
           Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sison,I.,
           Sodergren,B., Sonakke,T., Sparks,A., Stanley,H., Stone,H.,
           Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
           Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
           Uemari,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,
           Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
           Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
           Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
           Weinstein,G. and Gibbs,R.
:JOURNAL   Unpublished
:THORS     2 (bases 1 to 61954)
:REFERENCE Direct Submission
:THORS     Submitted (14-MAY-2002) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
           3 (bases 1 to 61954)
:REFERENCE Direct Submission
:THORS     Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
           On Jul 18, 2002 this sequence version replaced gi:20564251.
:JOURNAL   ----- Genome Center
           Center: Baylor College of Medicine
           Center code: BCM
           Web site: http://www.hgsc.bcm.tmc.edu/
           Contact: hgsc-help@bcm.tmc.edu
           ----- Project Information
           Center project name: GMBB
           Center clone name:
           ----- Summary Statistics

```

```

Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 78665 bases at least Q40
Consensus quality: 87559 bases at least Q30
Consensus quality: 93646 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2038: contig of 2038 bp in length
2039 2138: gap of unknown length
2139 4239: contig of 2101 bp in length
4240 4339: gap of unknown length
4340 6593: contig of 2254 bp in length
6594 6693: gap of unknown length
6694 8997: contig of 2304 bp in length
8998 9097: gap of unknown length
9098 11920: contig of 2823 bp in length
11921 12020: gap of unknown length
12021 14131: contig of 2111 bp in length
14132 14231: gap of unknown length
14232 16548: contig of 2317 bp in length
16549 16648: gap of unknown length
16649 18948: contig of 2300 bp in length
18949 19048: gap of unknown length
19049 21748: contig of 2600 bp in length
21749 21748: gap of unknown length
21749 24035: contig of 2287 bp in length
24036 24135: gap of unknown length
24136 26191: contig of 2056 bp in length
26192 26291: gap of unknown length
26292 28364: contig of 2073 bp in length
28365 28464: gap of unknown length
28465 31590: contig of 3126 bp in length
31591 31690: gap of unknown length
31691 33847: contig of 2157 bp in length
33848 33947: gap of unknown length
33948 36904: contig of 2957 bp in length
36905 37004: gap of unknown length
37005 40115: contig of 3111 bp in length
40116 40215: gap of unknown length
40216 44221: contig of 4006 bp in length
44222 44321: gap of unknown length
44322 47641: contig of 3320 bp in length
47642 47741: gap of unknown length
47742 50640: contig of 2899 bp in length
50641 50740: gap of unknown length
50741 54553: contig of 3813 bp in length
54554 54653: gap of unknown length
54654 58413: contig of 3760 bp in length
58414 58514: gap of unknown length
58514 61954: contig of 3441 bp in length.
-----
FEATURES
source
1..61954
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
2039..2138
/estimated_length=unknown
4240..4339
/estimated_length=unknown
6594..6693
/estimated_length=unknown
8998..9097
/estimated_length=unknown

```


AL442071.30:58993.	59120.	AL442071.30:59120.	59300.
AL442071.30:59514.	59668.	AL442071.30:61196.	61472.
AL442071.30:65820.	65948.	AL442071.30:66676.	66753.
AL442071.30:67368.	67529.	AL442071.30:68956.	69060.
AL442071.30:69697.	69792.	AL442071.30:71530.	71655.
AL442071.30:74120.	74272.	AL442071.30:76834.	76988.
AL442071.30:83480.	83606.	AL442071.30:84155.	84338.
AL442071.30:84741.	84910.	AL442071.30:84155.	84338.
AL442071.30:90301.	90439.	AL442071.30:89423.	89601.
AL442071.30:92363.	92525.	AL442071.30:91124.	91317.
AL442071.30:96518.	96625.	AL442071.30:96820.	96918.
AL442071.30:98103.	98255.	AL442071.30:98556.	98644.
AL442071.30:101128.	101263.	AL137853.12:4093.	4222.
AL137853.12:54400.	5633.	AL137853.12:7622.	7846.
AL137853.12:11990.	12508.	AL137853.12:13312.	13488.
AL137853.12:13756.	13879.	AL137853.12:15408.	15616.
AL137853.12:15975.	16292.	AL137853.12:22763.	22929.
AL137853.12:24645.	24801.	AL137853.12:27094.	27211.
AL137853.12:33076.	33138.	AL137853.12:33818.	34059.
AL137853.12:34859.	34988.	AL137853.12:36568.	36767.
AL137853.12:40049.	40544.	AL137853.12:41802.	43273.
AL137853.12:76989.	77150.	AL137853.12:77345.	77515.
AL137853.12:78643.	78809.	AL137853.12:82070.	82299.
AL137853.12:82576.	82902.	AL137853.12:83857.	83941.
AL137853.12:84442.	84683.	AL137853.12:85317.	85530.
AL137853.12:87150.	87269.	AL137853.12:87663.	87770.
AL137853.12:89029.	89247.	AL137853.12:90173.	90316.
AL137853.12:91097.	91105.	AL137853.12:92392.	92634.
AL137853.12:93889.	94115.	AL137853.12:95618.	95740.
AL137853.12:96389.	96440.	AL137853.12:96567.	96668.
AL137853.12:96762.	96935.	AL137853.12:97027.	97182.
AL137853.12:97329.	97526.	AL137853.12:98061.	98189.
AL137853.12:99247.	99450.	AL137853.12:102047.	102207.
AL137853.12:102307.	102513.	AL137853.12:102648.	102767.
AL137853.12:103209.	103378.	AL137853.12:105667.	105826.
AL137853.12:106772.	106965.	AL137853.12:107299.	107427.
AL137853.12:108322.	108465.	AL137853.12:109524.	109712.
AL137853.12:112985.	113105.	AL137853.12:113701.	113853.
AL365277.23:2001.	2011.	AL365277.23:3394.	3456.
AL365277.23:4645.	4799.	AL365277.23:6374.	6488.
AL365277.23:11377.	11494.	AL365277.23:22560.	22771.
AL365277.23:27363.	27492.	AL365277.23:28300.	29749.
/gene="MACE1"			
/locus_tag="RP11_69611.1-001"			
join(14552.	14823.	AL442071.30:5092.	5153.
AL442071.30:23924.	24013.	AL442071.30:25218.	26285.
AL442071.30:31846.	31938.	AL442071.30:35618.	36244.
AL442071.30:57099.	57211.	AL442071.30:57319.	57435.
AL442071.30:57976.	58095.	AL442071.30:58243.	58378.
AL442071.30:58993.	59120.	AL442071.30:59207.	59300.
AL442071.30:59514.	59668.	AL442071.30:61196.	61472.
AL442071.30:63680.	63948.	AL442071.30:66766.	66753.
AL442071.30:67368.	67529.	AL442071.30:68956.	69060.
AL442071.30:69697.	69792.	AL442071.30:71530.	71655.
AL442071.30:74120.	74272.	AL442071.30:76834.	76988.
AL442071.30:83480.	83606.	AL442071.30:84155.	84338.
AL442071.30:84741.	84910.	AL442071.30:84155.	84338.
AL442071.30:90301.	90439.	AL442071.30:91124.	91317.
AL442071.30:92363.	92525.	AL442071.30:93574.	93657.
AL442071.30:96518.	96625.	AL442071.30:96820.	96918.
AL442071.30:98103.	98255.	AL442071.30:98556.	98644.
AL442071.30:101128.	101263.	AL137853.12:4093.	4222.
AL137853.12:54400.			

AL137853.12:64442. 84683, AL137853.12:85317. 85523,
 AL137853.12:87150. 82769, AL137853.12:87663. 87770,
 AL137853.12:89029. 89247, AL137853.12:90173. 90376,
 AL137853.12:91097. 91105, AL137853.12:92392. 92623,
 AL137853.12:93889. 94115, AL137853.12:95618. 95740,
 AL137853.12:95839. 95940, AL137853.12:96567. 96688,
 AL137853.12:96762. 96935, AL137853.12:97027. 97182,
 AL137853.12:97329. 97526, AL137853.12:98061. 98189,
 AL137853.12:99277. 99450, AL137853.12:102047. 102202,
 AL137853.12:102307. 102513, AL137853.12:102648. 102767,
 AL137853.12:103209. 103378, AL137853.12:105667. 105826,
 AL137853.12:106771. 106965, AL137853.12:107299. 107427,
 AL137853.12:108322. 108465, AL137853.12:109524. 109712,
 AL137853.12:112985. 113105, AL137853.12:113701. 113853,
 AL1365277.23:23645. 2011, AL1365277.23:3394. 3556,
 AL1365277.23:4645. 4799, AL1365277.23:6374. 6448,
 AL1365277.23:11377. 11494, AL1365277.23:22560. 22771,
 AL1365277.23:27363. 27492, AL1365277.23:28300. 29749)
 /gene="WACF1"
 /locus_tag="RP11-695H1.1-001"
 /locus_tag="microtubule-actin crosslinking factor 1"
 /note="match: ESTs: Em:AA169873.1 Em:AF150358.1
 Em:AL138426.1 Em:ALG98142.1 Em:ALG99066.1 Em:ALH70538.1
 Em:ALH713492.1 Em:ALH734479.1 Em:AM043677.1 Em:AT032902.1
 Em:BE007711.1 Em:BE8396321.1 Em:BE67805.1 Em:BE871881.1
 Em:BE837593.1 Em:BE838157.1 Em:BE843659.1 Em:BE025979.1
 Em:BG571834.1 Em:BG993158.1 Em:BI600856.1 Em:BI769261.1
 Em:BM471359.1 Em:BM0014846.1 Em:BM009823.1 Em:BM128354.1
 Em:BO309642.1 Em:BO328711.1 Em:BO625288.1 Em:BX49007.1
 Em:BX493047.1 Em:BX508137.1 Em:BX646635.1 Em:CA399757.1
 Em:CB045418.1 Em:CB051751.1 Em:CB051752.1 Em:CB267890.1
 Em:CB850509.1 Em:CD106951.1 Em:CD643646.1 Em:CD676309.1
 Em:HL3461.1 Em:W61052.1 Em:Z21839.1
 match: CDNA5: Em:AB029290.1 Em:AL122291.1 Em:BE026707.1"
 join(14604. 14823, AL442071.30:5092. 5153,
 AL442071.30:33924. 24013, AL442071.30:28208. 28285,
 AL442071.30:31846. 31938, AL442071.30:56118. 56284,
 AL442071.30:57099. 57211, AL442071.30:57319. 57455,
 AL442071.30:57976. 58095, AL442071.30:58243. 58338,
 AL442071.30:58993. 59120, AL442071.30:59207. 59300,
 AL442071.30:59514. 59668, AL442071.30:61196. 61472,
 AL442071.30:65820. 65948, AL442071.30:66676. 66753,
 AL442071.30:67368. 67529, AL442071.30:68956. 69060,
 AL442071.30:69697. 69792, AL442071.30:71530. 71655,
 AL442071.30:74120. 74272, AL442071.30:76834. 76988,
 AL442071.30:83480. 83606, AL442071.30:84155. 84338,
 AL442071.30:84741. 84910, AL442071.30:89423. 89601,
 AL442071.30:90301. 90439, AL442071.30:91124. 91317,
 AL442071.30:92363. 92525, AL442071.30:95774. 95867,
 AL442071.30:96518. 96625, AL442071.30:96820. 96918,
 AL442071.30:98103. 98255, AL442071.30:98556. 98644,
 AL442071.30:101128. 101263, AL137853.12:14093. 14242,
 AL137853.12:5400. 5633, AL137853.12:7622. 7846,
 AL137853.12:11990. 12508, AL137853.12:13312. 13488,
 AL137853.12:13756. 13879, AL137853.12:15408. 15616,
 AL137853.12:15975. 16292, AL137853.12:22763. 22929,
 Query Match 18.1% Score 338.8; DB 5; Length 63187;
 Best Local Similarity 80.8%; Pred. No. 1.ee-96;
 Matches 409; Conservative 0; Mismatches 92; Indels 5; Gaps 1;

682 TATTTGGACAGAGCTCTTCCCTCTCAACCCAGGCTGAGATACAGTGCATGATCTCAG 741
 8564 TTTTGTGGATGAGAGCTCTTGTCTGTCTGCCAGGCTGAGAGCACTGAGCATGATCTCAG 8623
 742 TTCACTGACACCTCCACCTCCGGGGTTCAAGCAATCTCTGCTCAGCTCCTGAGTATG 801
 8624 CTCACCTGACACCTCCGCTCCGGGGTTCAAGCAATCTCTGCTCAGCTCCTGAGTATG 8683
 802 CTAGATTTACAGAAAGTGCACCTCCAGCTTCAAGCTAATTTTGTATTTTATAGAGATGC 861
 8684 CTGGGATTTACAGGTGTGTGTCACACCAACAGCTAATTTTGTATTTTATAGAGACGG 8743
 862 GCTTTTGGCATGTGGCCATGCTAGTCTGAGAACCCCGACCTTCAGGTGATCCGCTGCT 921
 8744 GGTTCACCATGTTGACACAGGCTGTCTGAAACCCCTGAACTCAGGTGATCTGCCATCT 8803
 922 TGGCCCTCCCAAGTGTGGGATTTGACGGCTGAGCCAGCCGACGCTGAGCTACTCC 981
 8804 CTGCCACCCAAAGTGTGGGATTTACAGGTCTGAGACACCGACCCGCGCTACTCTCTGCC 8863
 982 TTTAGTCTCTGGAAGACTGCGGCTC 1007
 8864 TCTATGAGCTCAATATTTTAGGTTTC 8889

.T 4
 1015/c 115710 bp DNA 1linear HTG 06-SEP-2001
 S Homo sapiens chromosome 1 clone RP4-648J17 map p34.1-34.3, 6
 NITION unordered pieces.
 SION AL139015
 ION AL139015
 ORDS AL139015.6 GI:13567858
 CE HTG; HTGS_PHASE1; HTGS_CANCELLED.
 GANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 Plumb,B.
 Direct Submission
 Submitted (05-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
 request: clonerequest@sanger.ac.uk
 On Apr 9, 2001 this sequence version replaced gi:9796296.
 ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 ----- Project Information
 Center project name: djf648J17
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 113662 bases at least Q40
 Consensus quality: 114216 bases at least Q30
 Insert size: 115210; sum-of-coverage
 Insert size: 112500; 15.0% error; agarose-fp
 Quality coverage: 6.13x in Q20 bases; sum-of-coverage
 coverage: 5.53x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 36330: contig of 36330 bp in length

FEATURES	source	1..115710
* 36331	36430: gap of 100 bp	
* 36431	47455: contig of 11025 bp in length	
* 47456	52799: gap of 100 bp	
* 47556	52799: contig of 5244 bp in length	
* 52800	52899: gap of 100 bp	
* 52900	60268: contig of 7369 bp in length	
* 60269	60368: gap of 100 bp	
* 60369	88988: contig of 26820 bp in length	
* 88989	89088: gap of 100 bp	
* 89089	115710: contig of 26622 bp in length.	
	Location/Qualifiers	
	1..115710	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="1"	
	/map="p34.1-34.3"	
	/clone="RP4-648J17"	
	/clone_id="RP4-4"	
misc_feature	1..36330	
	/note="assembly_fragment:01504"	
	fragment_chain:1	
misc_feature	36431..47455	
	/note="assembly_fragment:01208"	
	fragment_chain:1	
misc_feature	47556..52799	
	/note="assembly_fragment:00376"	
	fragment_chain:1	
misc_feature	52900..60268	
	/note="assembly_fragment:01176"	
	fragment_chain:1	
misc_feature	60369..88988	
	/note="assembly_fragment:00586"	
	fragment_chain:1	
misc_feature	89089..115710	
	/note="assembly_fragment:00793.0"	
ORIGIN		
Query Match	18.1%; Score 338.8; DB 12; Length 115710;	
Best Local Similarity	80.8%; Pred. No. 1.8e-96;	
Matches 409; Conservative 0; Mismatches 92; Indels 5; Gaps 1;		
QY	507 TTTTTCCTTTCTTTTCTTTTGGACAGAGCTCTGCTGCGCCAGGCTGAGTGC 566	
DB	3046 TTTATTTGTTGTTCTTTTGGATGAGATGAGTCTGCTTTAGCCAGGCTGAGTGC 2987	
QY	567 AGTGCAATGATCTGCGCCACGTGCAACCTCTGCTCCGGAATCAAGCATTCCTCTGC 626	
DB	2986 AGTGCAAGGTCTGAGCTCAGTCAAGCAACCTCCGCTCCAGGTTCAAGCATTCCTCTGC 2927	
QY	627 TCAGCTCCCAAGTACGCTGGATTTACAGGTGACAGCCACCAACCCAGC-----TTTTTT 681	
DB	2926 TCAGCTCCCAAGTACGCTGGATTTACAGGTGACAGCCACCAACCCAGC-----TTTTTT 2867	
QY	682 TATTTGGACAGAGCTCTTCCCTCTCAACCCAGGCTGAGATACAGTGCATGATCTCAG 741	
DB	2866 TTTTGTGGATGAGAGCTCTTGTCTGTCTGCCAGGCTGAGAGCACTGAGTATCTCAG 2807	
QY	742 TTCACTGACACCTCCACCTCCGGGGTTCAAGCAATTTCTGCTCAGTCTCTGAGTATG 801	
DB	2806 CTCACCTGACACCTCCGCTCCGGGGTTCAAGCAATTTCTGCTCAGTCTCTGAGTATCTCAG 2747	
QY	802 CTAGATTTACAGAAAGTGCACCTCCAGCTTCAAGCTAATTTTGTATTTTATAGAGATGC 861	
DB	2746 CTGGGATTTACAGGTGTGTGTCACACCAACAGCTAATTTTGTATTTTATAGAGACGG 2687	
QY	862 GCTTTTGGCATGTGGCCATGCTAGTCTGAGAACCCCGACCTTCAGGTGATCCGCTGCT 921	
DB	2686 GGTTCACCATGTTGACACAGGCTGTGCTGAAACCTCTGAATCTAGGTGATCTGCCATCT 2627	
QY	922 TGGCCCTCCCAAGTGTGGGATTTGACGGCTGAGCCAGCCGACGCTGAGCTACTCTCC 981	
DB	2626 CTGCCACCCAAAGTGTGGGATTTACAGGTCTGAGACACCGACCCGCGCTACTCTCTGCC 2567	

Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokweto, S., Ogih, M., Okumura, G., Otaguwa, N., Oyedro, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, B., Pu, L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G., Scheraga, S., Scott, G., Shen, H., Shim, C., Shooshtrati, N., Sisson, I., Sodergren, B., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansy, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Uman, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Kucherafatti, R., Weinstein, G., and Gibbs, R.

Unpublished

2 (bases 1 to 128829)

Worley, K.C.

Direct Submission

Submitted (12-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 128829)

Worley, K.C.

Direct Submission

Submitted (28-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 128829)

Worley, K.C.

Direct Submission

Submitted (25-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 128829)

Worley, K.C.

Direct Submission

Submitted (23-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

6 (bases 1 to 128829)

Worley, K.C.

Direct Submission

Submitted (15-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 25, 2002 this sequence version replaced gi:23343647.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2

clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

source

location/Qualifiers

1..128828

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="12"

/clone="RP11-15J12"

/complement(1..2000)

/note="overlaps bases 1..2000 of clone AC093016"

/function="clone overlap"

4..248

/rpt_family="AluSg/x"

499..791

/rpt_family="AluSx"

905..1203

/rpt_family="AluSg"

1499..1779

/rpt_family="AluSx"

1780..1809

/rpt_family="AluSg"

/rpt_family="AluSg"

1877..2102

/rpt_family="AluSg"

3349..3436

/rpt_family="AluSg"

3477..3513

/rpt_family="AluSg"

3529..3643

/rpt_family="AluSg"

4895..5183

/rpt_family="AluSg"

5184..5248

/rpt_family="AluSg"

/rpt_family="AluSg"

complement(5350..5655)

/rpt_family="AluSg"

5656..5683

/rpt_family="AluSg"

5686..5945

/rpt_family="AluSg"

5923..5925

/rpt_family="AluSg"

/function="low quality"

complement(6120..6421)

/rpt_family="AluSg"

6703..6755

/rpt_family="AluSg"

complement(6756..7068)

/rpt_family="AluSg"

7069..7196

/rpt_family="AluSg"

complement(7237..7544)

/rpt_family="AluSg"

7598..7738

/rpt_family="AluSg"

7739..8015

/rpt_family="AluSg"

8016..8056

/rpt_family="AluSg"

8057..8225

/rpt_family="AluSg"

8242..8264

/rpt_family="AluSg"

complement(8273..8385)

/rpt_family="AluSg"

```

repeat_region      complement(8389, .8698)
                    /rpt_family="AluXs"
repeat_region      complement(8927, .9215)
                    /rpt_family="AluSg"
repeat_region      9587, .9612
                    /rpt_family="("TTTTG)n"
repeat_region      complement(9613, .9898)
                    /rpt_family="AluSg"
repeat_region      9918, .10185
                    /rpt_family="AluXs"
repeat_region      10338, .10478
                    /rpt_family="AluJb"
repeat_region      10555, .10657
                    /rpt_family="L1MA8"

ary Match          17.9%; Score 335; DB 5; Length 128829;
at Local Similarity 81.8%; Pred. No. 3,1e-95;
ches 400; Conservative 0; Mismatches 85; Indels 4; Gaps 1;

502 TTTTCTTTTCTTTCTTTCTTTTCTTTTGTGAGACAGAGCTTGTCTGTGCCAGGCTG 561
    TTTTATTTATTTTATTTATTTATTTATTTGACAGAGTCTCACTGTCTCCAGGCTG 28862
    TTTTATTTATTTTATTTATTTATTTATTTGACAGAGTCTCACTGTCTCCAGGCTG 28862

562 AGTGCAGTGGCAGATCTCTGCCACTGCACCTCTGCTCCCGATTCAGAGATTC 621
    TTTTCTTTTCTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 28922
    AGTTCAGTGGCAGATCTGTGGCTCACTGCACCTCTGCTCTCAGGTTCAAGTATTC 28982

622 CTGCCTCAGCCTCCCAAGTGTGGATTACAGTGCAGCCACACCCAGCT---T 677
    CTGCCTCAGCCTCTGAGTGTGATTCAGATTCACAGCGGTGCACACCACTGGCTAAT 28982

678 TTTTATTTTGGACAGAGCTCTTCCCTCTGTACCCAGGCTGAGTACAGTGCATATC 737
    TTTTATTTTGGAGAGAGTCTGCTCTGTGTGCCAGGTTGAGAGTGCAGTGGCGTATC 29042

738 TCAGTTCACCTGGAGCCTCCACCTCCCGGGTTCAAGCAATTCCTCGTCACTGCTCA 797
    TTTTCTTTTCTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 29102
    TGGCTCACTGCACCTCCACCTCTCGGGTTCAAGGATTCCTCTGCTCAAGCTTCTCA 29162

798 GTAGTAGATTACAGAGTGCACCTCCACGTTCACTAATTTTGTATTTTATGATGAG 857
    GTAGCTGGAGACTACAGCGCTGGCCACACATGCCAGCTAATTTTGTATTTTATGATG 29162

858 ATGCGCTTTTGCATGTGTGCCATCTGTCAATCTGTGAAACCCCGGACCTCAGGTAT 917
    ATGGGGTTTACCATGTGTGCCAGGCTGTGTCTGAACTCTGTCACTCAAGTATCCGCC 29222

918 GCGTTGGCCTCCCAAGTCTGGGATTGACAGGCTGACAGCATTCGCCCGGCTAGATCA 977
    ACCCTAGCCTCCCAAGTCTGGGATTGACAGGCTGACAGCATTCGCCCGGCTAATTTT 29282

978 CTCCTTATG 986
    TTTTCTTTTATG 29342
    ATTTTATG 29350

AC026290      201508 bp      DNA      linear      HTG 07-JUL-2000
S      Homo sapiens chromosome 12 clone RP11-657111, WORKING DRAFT
NITION      SEQUENCE, 44 unordered pieces.
SION      AC026290
ON      AC026290.2 GI:8516761
RDS      HTG: HTGS PHASE1; HTGS_DRAFT.
E      Homo sapiens (human)
ANISM      Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homiidae; Homo.
          1 (bases 1 to 201508)
THORS      Waterston,R.H.
TLE      The sequence of Homo sapiens clone
URNAL      Unpublished

```

REFERENCE 2 (bases 1 to 201508)
AUTHORS Waterston, R. H.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2000) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis MO 63108, USA
COMMENT On Jun 14, 2000 this sequence version replaced gi:7272231c.

Genome Center	
Center: Washington University Genome Sequencing Center	
Center code: WUGSC	
Web site: http://genome.wustl.edu/gsc/index.shtml	
Project Information	
Center project name: H NH0657111	
Summary Statistics	
Sequencing vector: MJ3: 100%	
Sequencing vector: plasmid: 0%	
Chemistry: Dye-primed ET; 100% of reads	
Chemistry: Dye-terminator Big Dye; 0% of reads	
Assembly program: Phrap; version 0.990319	
Consensus quality: 183180 bases at least Q40	
Consensus quality: 188526 bases at least Q30	
Consensus quality: 190843 bases at least Q20	
Insert size: 223000; agarose-fp	
Insert size: 192700; sum-of-contigs	
Quality coverage: 3.87 in Q20 bases; agarose-fp	
Quality coverage: 3.45 in Q20 bases; sum-of-contigs	
NOTE: This is a 'working draft' sequence. It currently consists of 44 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	
1	1334: contig of 1334 bp in length
1335	1434: gap of unknown length
1435	2470: contig of 1036 bp in length
2471	2570: gap of unknown length
2571	3685: contig of 1115 bp in length
3686	3785: gap of unknown length
3786	5164: contig of 1379 bp in length
5165	5264: gap of unknown length
5265	7133: contig of 1869 bp in length
7134	7233: gap of unknown length
7234	8409: contig of 1176 bp in length
8410	8509: gap of unknown length
8510	10159: contig of 1650 bp in length
10160	10259: gap of unknown length
10260	11978: contig of 1719 bp in length
11979	12078: gap of unknown length
12079	13273: contig of 1345 bp in length
13424	13523: gap of unknown length
13524	14829: contig of 1306 bp in length
14830	14929: gap of unknown length
14930	17244: contig of 2315 bp in length
17245	17444: gap of unknown length
17445	19550: contig of 2606 bp in length
19551	20050: gap of unknown length
20051	23459: contig of 3409 bp in length
23460	23559: gap of unknown length
23560	26491: contig of 2932 bp in length
26492	26691: gap of unknown length
26692	29687: contig of 1096 bp in length
29688	29787: gap of unknown length
29788	32790: contig of 3003 bp in length
32791	32890: gap of unknown length
32891	35373: contig of 2483 bp in length
35374	35473: gap of unknown length
35474	38408: contig of 2935 bp in length
38409	38508: gap of unknown length
38509	41522: contig of 3014 bp in length


```

|||||
179801 ACCGACGCTCCCAAGTGTGAGATTTCAGCGGTGAGCCAGCGGCGCTATTTTT 179860
      978 CTCCTTTAG 986
      |||||
179861 ATTTTITTAG 179869

```

```

18
1320/c AC090320 191353 bp DNA linear HTG 09-MAY-2001
HTG Homo sapiens chromosome 18 clone RP11-892H1 map 18, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
SION AC090320
ON AC090320.4 GI:13518166
HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
E Homo sapiens (human)
ANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 191353)
BIRREN, B., LINTON, L., NUSBAUM, C. and LANDER, E.
HOMO SAPIENS CHROMOSOME 18, clone RP11-892H1
UNPUBLISHED
2 (bases 1 to 191353)
BIRREN, B., LINTON, L., NUSBAUM, C., LANDER, E., ALLEN, N., ANDERSON, S.,
BARRA, N., BASTIEN, V., BOGUSLAVSKY, L., BOUKHALTER, B., BROWN, A.,
CAMERATA, J., CAMPOLANO, A., CHOPEL, Y., COLANGELO, M., COLLINS, S.,
COLLYMORE, A., COOKE, P., DEARELLANO, K., DEWAR, K., DIAZ, J. S.,
DODGE, S., FARO, S., FERREIRA, P., FITZHUGH, W., GAGE, D., GALGAN, J.,
GARDYNA, S., GINDE, S., GOYETTE, M., GRAHAM, L., GRAND-PIERRE, N.,
HAGOS, B., HEAFORD, A., HORTON, L., HULME, W., ILLY, I., JOHNSON, R.,
JONES, C., KARATAS, A., LABOCQUE, K., LAMAZARES, R., LANDERS, T.,
LEHOCZKY, J., LEVINE, R., LIU, G., MACLEAN, C., MACDONALD, P.,
MARQUIS, N., MATTHEWS, C., MCCARTHY, M., MCKEAN, P., MCKERNAN, K.,
MCPEETERS, R., MEJDIRIM, J., MENEUS, J., MIHOVA, T., MINGA, V.,
MURPHY, T., NAYLOR, J., NGUYEN, C., NORBU, C., NORMAN, C. H.,
O'CONNOR, T., O'DONNELL, P., O'NEIL, D., OLIVER, J., PETERSON, K.,
PHUNKHANG, P., PIERRE, N., POLLARA, V., RAYMOND, C., RETTA, R.,
RIEBACK, M., RILEY, R., RISE, C., ROGOV, P., ROMAN, J., ROSETTI, M.,
ROY, A., SANTOE, R., SCHAUER, S., SCHUPBACK, R., SEAMAN, S., SEVERY, P.,
SOUNGUEZ, C., SPENCER, B., STRANGE-THOMANN, N., STOJANOVIC, N.,
STRAUSE, N., SUBRAMANIAN, A., TALAMAS, J., TESTA, S., THEODORE, J.,
TRAVERS, M., TRAVIS, N., TRIGILLO, J., VASSILIEV, H., VIEL, R., VO, A.,
WILSON, B., WU, X., WYMAN, D., YE, W. J., YOUNG, G., ZAHNOUN, J.,
ZEMBKE, L., ZIMMER, A. and ZODY, M.
Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 3, 2001 this sequence version replaced gi:13470217.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12591
Center clone name: 892_H1
----- Summary Statistics
Sequencing vector: P1asmid; n/a; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 187356 bases at least Q40
Consensus quality: 189415 bases at least Q30
Consensus quality: 190119 bases at least Q20
Insert size: 18300; agarose-fp
Insert size: 190453; sum-of-contigs
Quality coverage: 6.0 in Q20 bases; agarose-fp
Quality coverage: 5.7 in Q20 bases; sum-of-contigs

```

```

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 7710: contig of 7710 bp in length
* 7711 7810: gap of 100 bp
* 7811 8914: contig of 1104 bp in length
* 8915 9014: gap of 100 bp
* 9015 10614: contig of 1600 bp in length
* 10615 10714: gap of 100 bp
* 10715 14053: contig of 3339 bp in length
* 14054 14153: gap of 100 bp
* 14154 21518: contig of 7365 bp in length
* 21519 21619: gap of 100 bp
* 21619 75714: contig of 54096 bp in length
* 75715 75814: gap of 100 bp
* 75815 99969: contig of 24155 bp in length
* 99970 100069: gap of 100 bp
* 100070 130859: contig of 30790 bp in length
* 130860 130959: gap of 100 bp
* 130960 167261: contig of 36302 bp in length
* 167262 167262: gap of 100 bp
* 167262 191353: contig of 23992 bp in length.
Location/Qualifiers
1. 191353
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone.lib="RP11-892H1"
/clone.lib="RP11-892H1 Human Male BAC"
1. 7710
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
7711..7810
/estimated_length=100
7811..8914
/note="assembly_fragment"
8915..9014
/estimated_length=100
9015..10614
/note="assembly_fragment"
10615..10714
/estimated_length=100
10715..14053
/note="assembly_fragment"
14054..14153
/estimated_length=100
14154..21518
/note="assembly_fragment"
21519..21619
/estimated_length=100
21619..75714
/note="assembly_fragment"
75715..75814
/estimated_length=100
75815..99969
/note="assembly_fragment"
99970..100069
/estimated_length=100
100070..130859
/note="assembly_fragment"
130860..130959
/estimated_length=100
130960..167261
/note="assembly_fragment"

```



```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-104H15"
/clone_11b="RP11-104H15 Human Male BAC"
182..231
/notes="single clone coverage"
repeat_region complement(667..965)
/rpt_family="AluSc"
complement(1147..1227)
/rpt_family="L2"
2125..2168
repeat_region /rpt_family="GC_rich"
4585..4625
/rpt_family="(TG)n"
7353..7544
/rpt_family="MER20"
7552..7671
/rpt_family="GA-rich"
complement(7888..7995)
/rpt_family="MER5B"
complement(8099..8409)
/rpt_family="AluSq"
8421..8456
/rpt_family="AT_rich"
complement(8457..8753)
/rpt_family="AluSx"
8489..8496
/notes="<30 qual SINGL region"
8754..8787
/rpt_family="AT_rich"
complement(8788..8907)
/rpt_family="FLAM_C"
complement(8935..9026)
/rpt_family="MIR3"
9218..9273
/rpt_family="(TTCC)n"
complement(9274..9571)
/rpt_family="AluSx"
9469..9476
/notes="<30 qual SINGL region"
9482..9486
/notes="<30 qual SINGL region"
9501..9560
/notes="<30 qual SINGL region"
9587..9595
/notes="<30 qual SINGL region"
9618..9623
/notes="<30 qual SINGL region"
9777..9884
/rpt_family="MIR"
complement(9885..10211)
/rpt_family="AluSx"
10212..10228
/rpt_family="MIR"
10237..10378
/rpt_family="MER104"
10379..10680
/rpt_family="AluSx"
10681..10735
/rpt_family="AluSx"
10818..10967
/rpt_family="AT_rich"
10968..11265
/rpt_family="AluJo"
11266..11442
/rpt_family="AluSg"
11602..11902
/rpt_family="AluJo"
complement(11996..12227)
/rpt_family="L2"

```

```

repeat_region complement(12220..12452)
/rpt_family="L2"
repeat_region 12544..12842
/rpt_family="AluSc"
repeat_region complement(12886..13197)
/rpt_family="AluSx"
repeat_region complement(13275..13576)
/rpt_family="AluDb"
repeat_region 13683..14116
/rpt_family="MLTIG3"
repeat_region complement(14117..14424)
/rpt_family="AluY"
14425..14454
/rpt_family="MLTIG3"
repeat_region 14509..14632
/rpt_family="MIR"
repeat_region complement(14674..14962)
/rpt_family="AluSx"
14963..15131
/rpt_family="MLTIF"
repeat_region complement(15137..15248)
/rpt_family="MLTIF"
repeat_region complement(15249..15437)
/rpt_family="MIR"
15640..15933
/rpt_family="AluY"
complement(16136..16283)
/rpt_family="L2"

```

```

Query Match 17.8%; Score 332.6; DB 5; Length 184349;
Best Local Similarity 80.3%; Pred. No. 2.1e-94;
Matches 404; Conservative 0; Mismatches 94; Indels 5; Gaps 1;

QY 500 GCTTTTCTTTCTTTCTTTCTTTTGTGAGACAGAGCTTGTCTGTGCCAGGCT 559
DB 16366 GGGTTTCTTTTGTGATTTGTTTGTGAGACAGAGCTTGTCTGTGCCAGGCT 16425

QY 560 GGAGTGCAGTGCAGATCTCTGCCCACTGCAGCACTCTGCCCTCCGGATTCAAGCATTC 619
DB 16426 GGATTGCAGTGCAGATCTCTGCCCACTGCAGCACTCTGCCCTCCGGATTCAAGCATTC 16485

QY 620 TCTGCTCAGGCTCCCAAGTACGTGATTCAGAGTGCAGCCACCAACCCAGC---- 675
DB 16486 TCTGCTCAGGCTCCCAAGTACGTGATTCAGAGTGCAGCCACCAACCCAGC---- 16545

QY 676 -TTTTTTTATTTTGGACAGAGTCTTGCCTGTCACTCCAGCTGAGTACAGTGCATG 734
DB 16546 TTTTTTTTATTTTGGATGAGATCTGTCTGTCACTCCAGCTGAGTACAGTGCATG 16605

QY 735 ATCTCAGTTCACTGCAGCACTCCACCTCCCGGGTTCAAGCAATTCCTGCTCAAGTCTCC 794
DB 16606 ATCTCAGTTCACTGCAGCACTCCACCTCCCGGGTTTAAAGGATTCCTGCTCAAGTCTCC 16665

QY 795 TGAGTAGTACGATTACAGAGTGCACCTCCAGCTTCAAGTATTTTGTATTTTAACTA 854
DB 16666 AGAGTAGTACGATTACAGAGTGCACCTCCAGCTTCAAGTATTTTGTATTTTAACTA 16725

QY 855 GAGATGCGCTTTTGGCATGTGCGCATGTGATCTGGAATCCCGAGCTCAGGTATCCG 914
DB 16726 GAGATGCGGTTTTCACCATGTGTGCGCATGTGATCTGGAATCCCGAGCTCAGGTATCCG 16785

QY 915 CTGAGCTTTGGCTCCCAAGTGTCTGGATTTGACAGGCTGACGATCGCGCGGCTGAG 974
DB 16786 CCCAGCTTCCCTCCCTTAAAGTGTCTGGATTTGACAGGCTGACGATCGCGCGGCTGAG 16845

QY 975 CTATCCTTTAGTCTCTGGAAG 997
DB 16846 ATTAGTCTCTGATTTAGCAAG 16868

RESULT 10
CONSOLIDVI/c 187710 bp DNA linear PRI 04-OCT-2001
LOCUS

```

```

POSITION Human chromosome 14 DNA sequence BAC R-362122 of library RPC1-11
SESSION from chromosome 14 of Homo sapiens (Human), complete sequence.
AL115838
ID115838.6 GI:15982201
ORGANISM Homo sapiens (human)
REFERENCE Homo sapiens
AUTHORS Homo sapiens
TITLE Homo sapiens
JOURNAL Homo sapiens
GENOSCOPE Genoscope
DIRECT SUBMISSION Direct Submission
SUBMITTED (04-OCT-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
On Oct 5, 2001 this sequence version replaced gi:13872725.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : C-264121 (AC=AL163974)
Downstream BAC (overlapping the SP6 end) : R-63812
-----
Summary Statistics
Assembly Program: Phrap; version 2.0
Quality coverage: 6.92x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 : 1
1 - 9 : 53
10 - 19 : 407
20 - 29 : 1172
30 - 39 : 5394
40 - 49 : 11391
50 - 59 : 10988
60 - 69 : 11109
70 - 79 : 20167
80 - 89 : 49419
90 - 99 : 77609
-----
Percentage of bases with a quality value >= 40 : 96 %.
-----
Location/Qualifiers
1. 187710
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-362122"
/clone_11b="RPC1-11"
72840. 72970
/note="matching EMBL:R17876
RHdb:RH26378
dbSTS:STS8590
Identified using the e-PCR software (G. Schuler)"
87948. 88130
/note="matching EMBL:AA251131
RHdb:RH103649
dbSTS:STS71178
Identified using the e-PCR software (G. Schuler)"
111707. 111823
/note="matching EMBL:HI36689

```

Query Match	Best Local Similarity	82.6%	Pred. No. 2.1e-94;	Length 187710;
Matches 394;	Conservative	0;	Mismatches 79;	Indels 4; Gaps 1;
499	TGCTTTT	TTTTTTT	TTTTTTT	TGAGACAGAGTCTTGCTGCGCCAGCG
22393	TTCTGTTG	GGTAATTTCTTT	TTTTTTT	TGAGACAGAGTCTTGCTGCGCCAGCG
559	TGAGATG	CACTGCAAGATCTTGCCCACTGCACTTGCTCCCGGATTCAGAGCATT	618	
22333	TGAGATG	CACTGCAAGATCTTGCCCACTGCACTTGCTCCCGGATTCAGAGCATT	22274	
619	CTCTGCT	CAAGCTCCCAAGTACGAGTTTACAGATGACGCGCACACCGCAGC---	675	
22273	CTCTGCT	CAAGCTCCCAAGTACGAGTTTACAGATGACGCGCACACCGCAGCCTAA	22214	

134283 GGGTCTTTTGTGTTGATTGTTGTTTTTTAGACAGAGGCTCAGTCCTCCGCATCCAGGCT 134342

560 GGAGTGCAAGTGGCATATCTCTGCCCATCGAACCTCTGCTCCCGGATTCAGGCATTC 619

134343 GGAATTGAGGGCACAGACTTGGCTCATCTGCACCTTCTGCTTCCAGATTCAAAGTAATTC 134402

620 TCCGCTCAGACCCTCCCAAGTAGCTGGGAAATTACAGGTGCACGCGCACACCCAGC---- 675

134403 TCTGCTCAGCCTCTCGAGTAGTGGGAAATTACAGGGCCCCCGCACCATCCAGCTAAT 134462

676 ----TTTTTAAATTTTGGAGACAAGTCTTGGCCCTGTACCCAGGCTGAGTACAGTGC 731

134463 TTTTTTTTTTTTTTTTTGAAGTAGAGCTTGCTGTGTCAACCCAGGCTGAGTCAAGTGC 134522

732 ATGATCTCAGTTCACTCTCGACCTCTACACTCCCGGTTCAAGCAATTCCTCTGCTAGTC 791

134523 GCGATCTCGCTCATCTGCACCTCTGCTCTCTGGGTTAAGCAGATTCCTGCTTCAGCC 134582

792 TCCGAGTAGAGCTAGATTAAGAAGTGCACCTCCACGTTACAGCTAATTTTGTATTTTA 851

852 GTAGAGATGCGCTTTTGGCATGTTGGCCATGCTAGTGGAAACCCGGAAGCTCAGTGAT 911

134643 GTAAGACGGGGTTTCAACAATGTTGGCAAGGCTGTCTGGAACTCTGACCTCAGGTGAT 134702

912 CCGCTGGCTTGGCTCTCCCAAGTGTCTGGGATTCAGAGCGGTGAGCCATCCGCGCAGGCT 971

134703 CCACCCACCTCAGGCTCTCTAAGGCTGGGATTAAGGCGGTGACCCATCGCGCCGGCCA 134762

972 GAGCTACCTCTTGTGCTCTGGAAG 997

134763 GGGATTGGTGTCTTGATTTAGCAAG 134788

T 14
9554

DITTON Homo sapiens chromosome 17, clone RP11-542C16, complete sequence.

SION AC026954

RDS AC026954.14 GI:22773348

TE HTG.

ANISM Homo sapiens (human)

Homo sapiens

Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

1 (bases 1 to 101029)

Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-542C16
Unpublished
2 (bases 1 to 101029)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bedalov,F., Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G., Campiano,A., Castle,A., Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., Dextrallano,K., Dewar,K., Diaz,J.S., Dodge,S., Dominko,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Glade,S., Goyette,R., Graham,L., Grand-Pierre,N., Grant,G., Haeas,B., Heatford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lakocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N., McCarthy,M., McEwen,L., McGurt,A., McKernan,K., McPheters,R., Meldrum,J., Mihova,T., Miranda,C., Mlenka,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N., Plesni,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rotman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Teafay,S., Theodore,J., Tirrell,A., Trevers,M., Triggillo,J., Vasiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE	Direct Submission
JOURNAL	Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 101029)
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastein,V., Bloom,T., Boguslavsky,L., Bounghalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArrelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Katat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tsafaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (22-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	4 (bases 1 to 101029)
AUTHORS	Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastein,V., Bloom,T., Boguslavsky,L., Bounghalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArrelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Katat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tsafaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (10-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Sep 10, 2002 this sequence version replaced gi:22417404. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html
TITLE	Genome Center
JOURNAL	Center: Whitehead Institute/ MIT Center for Genome Research
COMMENT	Web site: http://www-seq.wi.mit.edu Contact: sequence.submissions@genome.wi.mit.edu
TITLE	Project Information
JOURNAL	Center project name: L8435
COMMENT	Center clone name: 542_C_16
FEATURES	Location/Qualifiers
SOURCE	1. 101029 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="17" /map="17" /clone="RP11-542C16" /clone_1lb="RPC1-11 Human Male BAC"

BLREN,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-199F11
Unpublished

(bases 1 to 121017)

BLREN,B., Lincoln,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavsky,I., Boukhalter,B., Brown,A.,
Camataia,T., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeKrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Dodge,S., Ferro,S., Ferreira,P., Fitzhigh,W., Gage,D., Galagan,J.,
Gardya,S., Glnde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Karatas,A., Karkas,A., Labocque,K., Lamazeres,R., Landers,T.,
Lachocky,J., Levine,R., Liu,G., Maclean,C., MacDonald,P.,
Marronchi,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Punhakhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schnpack,N., Seaman,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,C.N.,
Strausen,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Travers,M., Trevis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

RNAL Submitted (28-DEC-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

(bases 1 to 121017)

BLREN,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,I., Boukhalter,B.,
Camataia,T., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,A., Cooke,P., DeKrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Ferro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardya,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., MacDonald,P., Major,J., Matthews,C.,
Liu,G., Maclean,C., Macdonald,P., Maylor,J., Mlenga,V.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunhakhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schnpack,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

RNAL Submitted (03-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

(bases 1 to 121017)

BLREN,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,I., Boukhalter,B.,
Camataia,T., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,A., Cooke,P., DeKrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Ferro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardya,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunhakhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schnpack,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

RNAL Submitted (08-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Oct 8 2002 this sequence version replaced gi:22597626.

All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center -----

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information -----

Center project name: L11969

Center clone name: 199_F_11

Only the first 121 0 kilobases of this clone are being submitted
The remainder overlaps accession number AC007421 [WICGR project
L529].

FEATURES

source	1..121017
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="17"
	/map="17"
	/clone.lib="RP11-199P11"
unsure	/clone.lib="RPIC1-11 Human Male BAC"
	628..635
repeat_region	/note="<30 qual SNGL region"
	798..1141
unsure	/rpt_family="L2"
	1028..1033
repeat_region	/note="<30 qual SNGL region"
	complement(1142..1143)
repeat_region	/rpt_family="AluSx"
	1444..1969
repeat_region	/rpt_family="L2"
	complement(2081..2250)
repeat_region	/rpt_family="Charlie6"
	complement(2386..2681)
repeat_region	/rpt_family="AluY"
	2749..2907
repeat_region	/rpt_family="L1M4"
	2915..2949
repeat_region	/rpt_family="T(G)n"
	3005..3157
repeat_region	/rpt_family="AluDb"
	3158..3189
repeat_region	/rpt_family="(CA)n"
	3198..3685
repeat_region	/rpt_family="L1M4"
	3686..3714
repeat_region	/rpt_family="AT-rich"
	3718..3993
repeat_region	/rpt_family="AluSx"
	4431..4498
repeat_region	/rpt_family="HALL1"
	4513..4846
repeat_region	/rpt_family="HALL1"
	4928..5002
repeat_region	/rpt_family="AluS"
	5064..5093
repeat_region	/rpt_family="(TAA)n"
	complement(5101..5355)
repeat_region	/rpt_family="AluSg"
	5397..5579
repeat_region	/rpt_family="HALL1"
	complement(5712..6083)
repeat_region	/rpt_family="AluSg"
	6315..6617
repeat_region	/rpt_family="AluDb"
	7680..7840
repeat_region	/rpt_family="L1MC4a"
	7846..7961
repeat_region	/rpt_family="FLAM C"

```

unsure                                     7859. .7865
                                             /note=<30 qual SNGL region"
repeat_region                            complement(8097. .8400)
                                             /rpt_family="AluSg"
unsure                                    8289. .8293
                                             /note=<30 qual SNGL region"
unsure                                    8338. .8344
                                             /note=<30 qual SNGL region"
                                             complement(9099. .9212)
repeat_region                            /rpt_family="MIR"
                                             complement(9291. .9570)
repeat_region                            /rpt_family="AluJo"
                                             complement(9585. .9746)
repeat_region                            /rpt_family="L1PA13"
                                             9816. .10117
                                             /rpt_family="AluSx"
                                             complement(10792. .11012)
repeat_region                            /rpt_family="AluSg/x"
                                             11130. .11233
                                             /rpt_family="L2"
                                             complement(11429. .11726)
repeat_region                            /rpt_family="Aluub"
                                             complement(12660. .12962)
repeat_region                            /rpt_family="AluSx"
                                             13148. .13453
                                             /rpt_family="AluSp"
                                             13457. .13904
                                             /rpt_family="AluY"
                                             complement(14783. .14991)
repeat_region                            /rpt_family="MIR"
                                             complement(15683. .15986)
repeat_region                            /rpt_family="AluSx"
                                             16485. .16798
                                             /rpt_family="AluY"
                                             complement(18108. .18410)
repeat_region                            /rpt_family="AluSg"
                                             18938. .19057
                                             /rpt_family="AluJo"
                                             19058. .19361
repeat_region                            /rpt_family="AluSx"
                                             19377. .19673
repeat_region                            /rpt_family="Aluub"
                                             19681. .19998

```

Very Match	17.74;	Score 331;	DB 5;	Length 121017;
Bit Local Similarity	82.14;	Pred. No. 6e-94;		
Mismatches 395;	Conservative	0;	Mismatches 80;	Indels 6;
			Gaps	1;

502 TTTTCTTTCTTTCTTTCTTTCTTTTGGAGACAGACTCTGCTCTGTGCCAGGCTGG 561
 34111 TCTTCTTTTTTTTTTTTTTTTTTTTTTTTGGACAAATTCCTCACTGTCTCCAGGCTGG 34052
 562 AGTCAGATGGCATGATCTGTGCCCATCTGCAACTCTGCTCCGAGATTCAAGGATTCTC 621
 34051 AGTCATTGGCATATCTCACTCACTGCAACTCTGTGCTCCGAGCTCAAGATTCTT 33992
 622 CTGCTCAGCCTCCCAAGTACGTGGATTACAGGTGACGCGCACCAACCCAGC----- 675
 33991 GTGCTCAGCCTCCCAAGTACGTGGATTACAGGCAATACACACATAGCCAGTATTAT 33932
 676 TTTTCTTTATTTTGGAGACAGACTCTTGCCCTGTGACCCAGGCTGGAGTACAGTGGCATGA 735
 33931 TTTTATTTTATTTTGGACAGAGTCTCTCTCAAGTTGCGCAAGGCTGAGGTGACAGTGGCGCA 33872
 736 TCTCAGTTCACATGGGCACTCCACTCCCGGGTTCAACCAATTCCTCGTCCGCTCAGTCTCCT 795
 33871 TCTCAGTTCACATGGCACTCCACTCCCGGGTTCAACCAATTCCTCGTCCGCTCAGCTCCT 33812
 796 GAGTAGCTAGGATTACAGAAATGTGACCTCCACGTTCACTAATTTTGTATTTTTATAGTAG 855
 33811 GAGTAGCTGGGATTACAGGCGCTGTGGCACAACACCCGGCTAATTTTGTATTTTTATAGTAG 33755
 856 AGATGCGCTTTTGGCAATGTTGGCAATGCTAGTCTGGAACCCCGGACCTCAGTGATTCGC 915

Db 33751 AGATGGGGTTTCAACATGTTGGCCAAAGCTGCTCGAATCTCGATCTTGAAGTATCCAC 336922

QY 916 TGGCCTTGGGCTCTCCCAAAGTGTCTGGGATTGCAAGGCTGAGCCATGCGGCGCAAGGCTGAGC 975

Db 33691 CTGCTCTGGGCTCTCCCAAAGTGCTGGGATTACAAGGCGTGAAGCCACCGCACCTCAGTCTCCCC 33632

QY 976 T 976

Db 33631 T 33631

Search completed: June 22, 2006, 00:39:18
Job time : 10376 secs

This page Blank (unpsd)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

nucleic - nucleic search, using sw model

on: June 21, 2006, 21:42:46 ; Search time 9227 Seconds
(without alignments)
11332.952 Million cell updates/sec

US-10-502-279-26

1870
1 atctgtcctagaagtaac.....catcccgatcctcgtag 1870

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

48236798 seqs, 27959665780 residues

number of hits satisfying chosen parameters: 96473596

DB seq length: 0

DB seq length: 200000000

processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

base :

EST.*
1: gb_est1.*
2: gb_est3.*
3: gb_est4.*
4: gb_est5.*
5: gb_est6.*
6: gb_est7.*
7: gb_est8.*
8: gb_est9.*
9: gb_est10.*
10: gb_est11.*
11: gb_est12.*
12: gb_est13.*
13: gb_est14.*
14: gb_est15.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Hit	Score	Query Match	Length	DB ID	Description
1	365	19.5	365	1	AA837817
2	324.4	17.3	475	4	BX505056
3	309	16.5	484	2	BG548177
4	306.8	16.4	3143	6	BSM805999
5	305.8	16.4	648	4	CR217138
6	305	16.3	520	8	CR556374
7	304.8	16.3	648	14	DU636687
8	304.2	16.3	679	14	AG111125
9	303.6	16.2	4640	6	CR857699
10	303	16.2	515	9	DB062128
11	299.6	16.0	742	5	CP127609
12	298.8	16.0	620	13	CZ450138
13	298.8	16.0	743	3	B0629138
14	298.6	15.9	720	3	CR871743
15	297.2	15.9	537	2	BF876683
16	296.6	15.9	583	3	BQ007997
17	295.8	15.8	588	3	DB258133
18	295.6	15.8	573	9	DA180845
19	295.2	15.8	678	14	AG118253

20	294.8	15.8	608	7	AW970856
21	294	15.7	560	9	DA780305
22	294	15.7	8535	6	CR933630
23	293.6	15.7	541	4	BX508653
24	293.2	15.7	489	7	BF673854
25	292.4	15.6	514	13	CZ456835
26	292.4	15.6	1084	3	BM806108
27	292.4	15.6	1467	6	BC039500
28	292.2	15.6	559	9	DB196716
29	292.2	15.6	799	13	CZ461615
30	292.2	15.6	884	13	CZ459857
31	291.8	15.6	543	9	DB343878
32	291.8	15.6	635	11	AQ338583
33	291.8	15.6	648	14	AG128024
34	291.6	15.6	496	8	CN315352
35	291.4	15.6	550	9	DB102123
36	291.4	15.6	771	1	AA984258
37	290.6	15.5	680	3	BQ181560
38	288.8	15.4	508	13	CZ456641
39	288.6	15.4	517	4	CA434698
40	288.4	15.4	579	4	BX646636
41	288	15.4	612	7	AW979073
42	287.8	15.4	3820	6	BSM803830
43	287.8	15.4	3820	6	BSM803831
44	287.2	15.4	718	14	CR958304
45	287	15.3	498	2	B1551454

ALIGNMENTS

RESULT 1
AA837817
LOCUS
DEFINITION
AA837817
365 bp mRNA linear EST 18-MAR-1998
0639H03.81 NCI CGAP Pr25 Homo sapiens CDNA clone IMAGE:1410965
similar to contains Alu repetitive element; contains element MBR36
repetitive element ; mRNA sequence.

ACCESSION
AA837817
VERSION
AA837817.1 GI:2913474
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE
1 (bases 1 to 365)
NCI CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright,
Ph.D.

CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdip/image/image.html
Insert length: 1141 Std Error: 0.00
Seq primer: -40m13 fwd. RT from Amersham
High quality sequence stop: 333.
Location/Qualifiers
1. 365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1410965"
/isue_type="epithelium (cell line)"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP Pr25"
/note="Organ: prostate; Vector: Bluescript SK-, Site_1"

FEATURES

source

1991 S. Sugiyama, T. Saito, K. Isono, Y. Irie, R. Kuehida, N. Yoneyama, T. Otsuka, R. Kanda, K. Yokoi, T. Kondo, H. Waga, S. M. Murakawa, K. Ishida, S. Ishibashi, T. Takahashi, Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isegai, T. and Sugano, S.

1992 Identification and Characterization of Putative Alternative Promoters of Human Genes

1993 Genome Res. 16 (1), 55-65 (2006)

1994 Contact: Takao Isegai

1995 Helix Research Institute

26-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: fujii-cdn@nifty.com

1996 NEDO human cDNA project (New Energy and Industrial Technology Developmental Organization (NEDO), cDNA library construction; Helix Research Institute for Biotechnology (RIB) and Biotechnology Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing; RAB.

1997 Location/Qualifiers

1. .515

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BEST14002179"

/issue_type="testis"

/clone_lib="BEST14"

/note="Vector: pMER18F13"

1998

1999

2000

2001

2002

2003

2004

2005

2006

2007

2008

2009

2010

2011

2012

2013

2014

2015

2016

2017

2018

2019

2020

2021

2022

RESULT 11

CF127609

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1. .742

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGS:3053058"

/issue_type="Chondrosarcoma Lung Metastasis cell lines"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH MGC 214"

/note="Organ: Lung; Vector: pYX-Asc; Site: 1; Ecor I; Site: 2; Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction. Ligated with Ecor I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

1998

1999

2000

2001

2002

2003

2004

2005

2006

2007

2008

2009

2010

2011

2012

2013

2014

2015

2016

2017

2018

2019

2020

2021

2022


```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FLO-bdh-1-10-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FLO"
/notes="Organ: Chondrosarcoma; Vector: pT73-Pac
(Pharmacia) with a modified polylinker; Site 1: Ecor I;
Site 2: Not I; NCI CGAP FLO is a cDNA library derived from
a pool of mRNA obtained from 4 cell lines from grade III
chondrosarcoma tissues. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GAGGTCGCTG. The cell line
was provided by Dr James Martin from University of Iowa.
TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG LIB=UI-H-FLO
TAG_SEQ=GAGGTCGCTG"

```

```

% Match 16.0%; Score 298.8; DB 3; Length 743;
% Local Similarity 76.1%; Pred. No. 3.8e-37;
% Conservative 0; Mismatches 107; Indels 15; Gaps 1;

```

```

512 TTTCTTTTCTTTTGTGAGACAGAGCTTGTCTGTGCTCCGCGAGCTGAGTGAAGTGG 571
1 TTTTCTTTTCTTTTGTGAGACAGAGCTTGTCTGTGCTCCGCGAGCTGAGTGAAGTGA 60
572 CAGATCTCTGCGACCTGCACTCTGCTCCGCGAGTCAAGCATCTCTGCTGAC 631
61 CGCATCTCTGCTCAGTCAAGCTCCGCTCTGCGGTCAATGCTCTCTGCTGTC 120
632 CTCCCAAGTAGCTGGAGTACAGGTGACGCGCACACACCGACG-----T 676
121 CTCCCAAGTAGCTGGAGTACAGGTGACGCGCACACCTACAGCAATTTTGGCATTT 180
677 TTTTATTTTGTGAGACAGAGCTTGTGCTGTCACTCCGCGAGCTGAGTGAAGTGA 736
181 TTTATTTTGTGAGTCAAGGTCTCGCTATATCCCGAGCTGAGTGAAGTGAAGTGA 240
737 CTCAATCTGAGTCACTGCACTCCGCGGTCAAGCAATTTCTGCTCAAGTCTCTG 736
241 CTGGGCTCAGTCACTGCACTCCGCGGTCAAGCAATTTCTGCTCAAGTCTCTG 300
797 AGTAGTAGAGTACAGAGTGAACCTCAAGTCAAGTCAATTTTGTATTTTATGAGA 856
301 AGTAGTAGAGTACAGAGTGAACCTCAAGTCAAGTCAATTTTGTATTTTATGAGA 360
857 GATGCGCTTTTGGCATTTGGCCATGCTAGTGAACCCCGAGCTCAGATGATCCGCT 916
361 GATGCGGTTTGGCATTTGGCCATGCTAGTGAACCTCAGATGATGATGATGATGAT 420
917 GGGCTTGGCTCCCAAGTGTGAGATGAGGCTGAGCAATGCGCGAGCTGAGCT 976
421 TGCGCTGGCTCCCAAGTGTGAGATGAGGCTGAGCAATGCGCGAGCTGAGCT 480
977 ACTGCTTGTAGTCTGGAAGAGCTGCGGCT 1006
481 TTGGGTTTTTTTGTATGATGAGATGAGGCTT 510

```

```

117 14 720 bp mRNA linear EST 03-FEB-2005
:743/c CX871743
HESCA_60.G05.G1.A037 NIH_MGC_262 Homo sapiens cDNA clone
IMAGE:7486284 5', mRNA sequence.

```

```

ACCESSION CX871743
VERSION GI:58554917
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 720)
NIH-MGC http://imgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
REFERENCE
AUTHORS Unpublished (1999)
TITLE Contact: Daniela S. Gerhard, Ph.D.
JOURNAL Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@bld-31.nih.gov
Tissue Procurement: Breasden, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
University of Georgia
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
Plate: L1A15813 row: n column: 10
Seq primer: JENREV (CAGGAACAGCTATGACC)
High quality sequence stop: 720.
Location/Qualifiers
1..720
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7486284"
/sex="male"
/tissue_type="embryonic stem"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/notes="Vector: DEKpress-1; Site 1: NotI; Site 2: EcoRV;
RNA obtained from human embryonic stem cells isolated from
the inner cell mass of blastocyst stage embryos and
differentiated to an early neural progenitor cell type.
Cell line id and NIH Registry designation is BG01.
Positive for Nestin and Musashi expression. Passage number
18. cDNA primed using oligo-dT primer:
5'-pACTGTTTCTAGATGCGGAGGCGGCTT25-3' and cloned into
the EcoRV/NotI sites of DEKpress-1. This primary library
is non-normalized (normalized primary library is
NIH MGC 259). It was constructed by Express Genomics
(Federick, MD). Sequence ends have been trimmed to
exclude vector and regions below phred quality 16. Note:
this is a Mammalian Gene Collection library."

```

```

ORIGIN
Query Match 16.0%; Score 298.6; DB 9; Length 720;
Best Local Similarity 79.1%; Pred. No. 4.1e-37;
Matches 406; Conservative 0; Mismatches 99; Indels 8; Gaps 4;
QY 484 AGATTAGAGTCTGCTGCTTTTCTTTTCTTTTCTTTTGTGAGACAGAGTCTTG 543
DB 634 AGCTGGAGACTACGTTTGTGTTGTTGTTTGTGTTTGTGTTTGTGAGACAGAGTCT-- 577
QY 544 CTCTGTGCGCGAGGCTGAGTGAAGTATCTCTGCCACAGCACTCTGCTCC 603
DB 576 CACTGTACCCAGAGCTGAGAGTGTGTGTATCTCGGTACCTCAACCTCTGCTCC 517
QY 604 CGGATTCAAGCAATCTCTGCTCAGCTCCCAAGTATGAGATTAAGTGAAGTGAAGCC 663
DB 516 CGGGTTCAAGTATTTCTCGCTCAGCTCCGAGTATGAGATTAAGTGAAGTGAAGCC 457
QY 664 ACCACACCAAGC---TTTTTTTATTTTGTGAGACAGAGTCTTGGCTCTGACCAAGGCTGG 720

```

456 ACCATGCGCTGGCTAAATTTTCTTTTCTTTTGTAGAGACAGAGCTCGCTCTGTGCGCCAGGCTGG 397

721 AGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTC 780

396 AGTGCAGGTGGTGGGCTTTGGCTCACTGCAACCTCCGCTCCAGGTTTACGCAATTTCTC 337

781 CTGGCCCACTCTCCGAGTACGAGTACGAGTATTAAGAAAGTGCACCTCCACGTTAGCTAA-TT 839

336 CTGCTCAGCTCTCTGAGTACGAGTGGAGCTAGGAGCGCCGCCGACCAAGCCCACTAAATTT 277

840 TTTGTATTTTGTAGTACAGATGCGCTTTTGGCCATGTTTGCCCATGTAGTCTGTAAGAACCCGG 899

276 TTTGTATTTTGTAGTACAGATGAGGGGTTTCACTGTGTTAGCAAGATGATCTGATCTCTC 217

900 ACTCTGAGTATTCGCTGGCTTTGGCTTCCCAATGCTGGGATTTGAGGCGGTAGGCAT 959

216 ACTCTT--GTGATCCACCGCCCTCGGCTCTTAAAGTGTGGGATTTACAGGCGTATCAT 159

960 CGCGCAGGCGCTGAGCTACCTCTTATGTTCTTG 992

158 CGCATCCGCGAGCACTTCTCCCTTCTGTCTATG 126

537 bp mRNA linear EST 17-JAN-2001

BF876683

0V0-ET0149-131100-496-b06 ET0149 Homo sapiens cDNA, mRNA sequence.

BF876683

BF876683.1 GI:1226722

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 537)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zagzo,M.A., Bordin,S., Costa,F.P., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.V.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=0V0et2=QV0-ET0149-131100-496-b06et3=2000-11-13et4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 13

High quality sequence stop: 534.

Location/Qualifiers

1. 537

source

ORIGIN

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match	Score	DB 2;	Length
15.98;	297.2;		5377;
Best local similarity:	78.48;	Prod NO. 7	5a-37.

Matches 373; Conservative 0; Mismatches 93; Indels 10; Gaps 1.

519 TCTTTTGTGAGACAGAGTCTTGCTCTGTCCGCCAGGCTGGAGTGCAGTGGCATGATC 578

Db 521 TTTT TTTT TGAGACAGAGTTCACTCTGTTGCCCAAGCCGGAGTGCAGTGGTGATC 462

579 TCTGCCACTGCACCTCTGCCTCCCGATTCAAGGATTCTCCTGCCTCAGCCTCCCA 638

Db 461 GCAGCTCACTGCAACCTCCGCCCTCCAGATTCCAGCAATTCTTCCACCCTCAGCCTCCTGA 402

639 GTAGCTGGATTACAGGTGCACGCCACCACCCAGCT-----TTTTTATTTTG 688

Db 401 GTAGCTAGGATTACAGACGTGCACCAACCATGCCAGCTAATTTATTTATTTATTTT 342

689 GAGACAGAGTCTTGCCCTGTCAACCAGGCTGGAGTACAGTGGCATGATCTCAGTTCACCTG 748

Db 341 GAGACGGAGTTTCACTCTGTCACCAGGCTGGAGTGCAGTGGCACTGTCTCGTTACTG 282

749 CGACCTCCACCTCCCGGTTCAAGCAATTCTCCTGCCCTCAGTCTCCCTGAGTAGCTAGGAT 808

Db 281 CAACCTCTGCCTCCCAAGGTTCAAGCAATTCTCCTGCCTGGCCCTCCCAAGTAGCTGGAC 222

809 TACAGAAGTGCACCTCCACGTTACGCTAATTTTGTATTTTCTAGTAGAGATGCCCTTTTG 868

Db 221 TACAGGCGTGTGCCAACACGCCCAGCTAATTTTGTATTTTGTAGTAGAGACAGGGTTTG 162

869 CCATGTTGGCCATGCTAGTCTGGAAACCCCGGACCTCAGGTGATCCGCTGGCTTTGGCCCTC 928

Db 161 CCATGTTGCCTAGGCTGTTCTCAAACCTCCTGACCTCAGGTGATCCACCCTGGCTTC 102

929 CCAAGTGTGGGATTGCAGGCGTGAGCCATCGCCCGAGGCTGAGCTACTCTT 984

Search completed: June 22, 2006, 01:59:31
Job time : 9233 secs

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/def_cstage="Adult"  
/clone_id="E70149"  
/note="Organ: lung_tumor; Vector: puc18; Site:1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORSTS PCR (U.S. Letters Patent application  
NO. 196,716 - Ludwig Institute for Cancer Research)
```


975 CTACCTCCTTACTCTCTGGAAGAAGTCGGCTCAGAGAAATCAACGCT 1022
5517 TTGCTTCTTTTCTTTTCTGAGACAGTCTCTTCTGTCAACCCAGGCT 5564

1. IT 4
 2. 9-949-002-825/c
 3. Sequence 825, Application US/09949002
 4. Client No. 6900016
 5. SERIAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION

FILE REFERENCE: CL000790

IRRENT APPLICATION NUMBER: US/09/949,002

CURRENT FILING DATE: 2000-01-28
 PTO APPLICATION NUMBER: 60/231

PRIOR APPLICATION NUMBER: 60/231,401
 PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 10823

()FTWARE: FastSEQ for windows Version 4.0

Q ID NO 825

LENGTH: 45819

TYPE: DNA

ORGANISM: Human
(1)-949-002-825

64-70207-13

Easy Match

```

Query Match      17.3%; Score 323.6; DB 3; Length 45819;
Hit Local Similarity 82.9%; Pred. No. 5.9e-79;
Matches 383; Conservative 0; Mismatches 74; Indels 5; Gaps 1;

```

517 TTCTTTTGGAGCAGAGTCTGCTCTGTGCCCCAGGCTGGAGTGCAGTGCATGA 576

23998 TTTT TTTT TTTGAGATGGAGTCTCACTCTGTCCCTAGGGTGGAGTGCAGTGGCCAA 23935

577 TCTCTGCCCACTGCACCTCTGCCCTCCCGATTCAAGGATTCTCCTGCCTCAGCCTCCC 636

23938 TCTCAGCTCACTGCAACCTCCGCCCTCCCGGTTCAAGAGATTCTCCTGCTCAGCCTCCC 23875

637 AAGTAGCTGGGATTACAGGTGCACGCCACCAACCAGC-----TTTTTTATTTTGAG 691

23878 AAGTAGCTGGACTACAAGCACATGCCACTGCCGCCAGCTAATTTTTTTTTTTGAG 23819

692 ACAGAGTCTTGCCCTGTCAACCAGGCTGAGTACAGTGGCATGATCTCAGTTCACCTGCGA 751

23818 ATGAGTCTACTCTGTGCGCTAGGGTGAGTGCAGTGGCGCAATCTCAGCTCACTGCA 23759

752 CCTCCACCTCCGGGTTCAAGCAATTCTCCGCTCAGTCTCCTGAGTAGCTAGGATTAC 811

23758 CCTCCGCTCCGAGTCAAGTATCTCTGCTCAGCCTCCTGAGTAGCTGGATTAC 23699

812 AGAGTGCACCTCCACGTTACGTAATTTTGTATTTTGTAGAGATGGCTTTGCCA 871

23698 AGGTGCCCAACCAACGCGCACTAATTTGTATTTAGTAGAGACGGTGTGCA 23639

872 TGTGGCCATGCTAGTCTGGAACCCCGACCTCAGGTGATCCGCTGGCCTTGCCCTCCA 931

23638 TGTGGCAGGCTGGTCTGGCACTAGACCTCAGGTGATCCACCCGCTCGGCTCCA 23579

932 AAGTCTGGATTGCAGGCGTGAGCCATCGCCAGGCTGA 973

23578 AAGTCTGGATTACAGCATGAGCCAGCGGCCGCCAGA 23537

1.5

91-345-882-1/c

Application US/09345882

FILE NO. 6399373
LIBRARY INFORMATION

APPLICANT: Bouquet

TITLE OF INVENTION: A NUCLEIC ACID SEQUENCE

TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.

TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.

FILE REFERENCE:	GENSET..031A
CURRENT APPLICATION NUMBER:	US/09/345,882
CURRENT FILING DATE:	1999-06-30
PRIOR APPLICATION NUMBER:	US 60/091,315
PRIOR FILING DATE:	1998-06-30
PRIOR APPLICATION NUMBER:	US 60/111,909
PRIOR FILING DATE:	1998-12-10
NUMBER OF SEQ ID NOS:	140
SOFTWARE:	Patent .pm
SEQ ID NO 1	
LENGTH:	162450
TYPE:	DNA
ORGANISM:	Homo sapiens
FEATURE:	
NAME/KEY:	allele
LOCATION:	72794
OTHER INFORMATION:	5-124-273 : polymorphic base A or G
FEATURE:	
NAME/KEY:	allele
LOCATION:	88073
OTHER INFORMATION:	5-127-261 : polymorphic base A or C
FEATURE:	
NAME/KEY:	allele
LOCATION:	90842
OTHER INFORMATION:	99-1437-325 : polymorphic base A or G
FEATURE:	
NAME/KEY:	allele
LOCATION:	93714
OTHER INFORMATION:	5-128-60 : polymorphic base deletion of GT
FEATURE:	
NAME/KEY:	allele
LOCATION:	97122
OTHER INFORMATION:	99-1442-224 : polymorphic base G or T
FEATURE:	
NAME/KEY:	allele
LOCATION:	97152
OTHER INFORMATION:	5-129-144 : polymorphic base deletion of T
FEATURE:	
NAME/KEY:	allele
LOCATION:	99098
OTHER INFORMATION:	5-130-257 : polymorphic base A or G
FEATURE:	
NAME/KEY:	allele
LOCATION:	99117
OTHER INFORMATION:	5-130-276 : polymorphic base A or G
FEATURE:	
NAME/KEY:	allele
LOCATION:	103806
OTHER INFORMATION:	5-131-395 : polymorphic base A or T
FEATURE:	
NAME/KEY:	allele
LOCATION:	106940
OTHER INFORMATION:	5-133-375 : polymorphic base insertion of A
FEATURE:	
NAME/KEY:	allele
LOCATION:	108106
OTHER INFORMATION:	5-135-155 : polymorphic base insertion of A
FEATURE:	
NAME/KEY:	allele
LOCATION:	108149
OTHER INFORMATION:	5-135-198 : polymorphic base insertion of GTTTT
FEATURE:	
NAME/KEY:	allele
LOCATION:	108308
OTHER INFORMATION:	5-135-357 : polymorphic base A or G
FEATURE:	
NAME/KEY:	allele
LOCATION:	108471
OTHER INFORMATION:	5-136-174 : polymorphic base C or T
FEATURE:	
NAME/KEY:	allele
LOCATION:	134134
OTHER INFORMATION:	5-140-120 : polymorphic base C or T

```
FEATURE:  
NAME/KEY: allele  
.CCATION: 134362  
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A  
FEATURE:  
NAME/KEY: allele  
.CCATION: 134374  
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA  
FEATURE:  
NAME/KEY: allele  
.CCATION: 146328  
OTHER INFORMATION: 5-143-84 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
.CCATION: 146345  
OTHER INFORMATION: 5-143-101 : polymorphic base A or C  
FEATURE:  
NAME/KEY: allele  
.CCATION: 150329  
OTHER INFORMATION: 5-145-24 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
.CCATION: 160031  
OTHER INFORMATION: 5-148-352 : polymorphic base G or T  
FEATURE:  
NAME/KEY: allele  
.CCATION: 72771..72817  
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30  
FEATURE:  
NAME/KEY: allele  
.CCATION: 72771..72817  
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30  
FEATURE:  
NAME/KEY: allele  
.CCATION: 72771..72817  
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID31  
FEATURE:  
NAME/KEY: allele  
.CCATION: 88050..88096  
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31  
FEATURE:  
NAME/KEY: allele  
.CCATION: 88050..88096  
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31  
FEATURE:  
NAME/KEY: allele  
.CCATION: 90819..90865  
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49  
FEATURE:  
NAME/KEY: allele  
.CCATION: 90819..90865  
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49  
FEATURE:  
NAME/KEY: allele  
.CCATION: 90819..90865  
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70  
FEATURE:  
NAME/KEY: allele  
.CCATION: 93690..93736  
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32  
FEATURE:  
NAME/KEY: allele  
.CCATION: 93690..93736  
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32  
FEATURE:  
NAME/KEY: allele  
.CCATION: 93690..93736  
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53  
FEATURE:  
NAME/KEY: allele  
.CCATION: 97099..97145  
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50  
FEATURE:  
NAME/KEY: allele  
.CCATION: 97099..97145  
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50  
FEATURE:  
NAME/KEY: allele  
.CCATION: 97130..97177  
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33  
FEATURE:  
NAME/KEY: allele  
.CCATION: 97130..97177  
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54  
FEATURE:
```

DB	Query Match	17.3% Best Local Similarity	Score 323.6; 81.2% 401; Matches	DB 3; Pred. No. 1.1e-78; 0; Conservative	Length 162450; Indels 89; Mismatches 89;	Gap 4; Gaps 2;
QY	488	TTGAGCTGCTGCTGCTTTT	TTT	TTT	TTT	TTT
DB	102008	TAAAGAGAGAGTAATCTTT	TTT	TTT	TTT	TTT
QY	548	GTCGCCAGGCGTGAAGTGA	GATGATCTCTG	CCCACTGCAACTCTG	CTCCCGGA	607
DB	101948	GTCGCCAGGCGTGAAGTGA	GATGATCTCTG	CCCACTGCAACTCTG	CTCCCGGA	101889
QY	608	TTCAAGCATTTCTCTGCTCA	CGCTCCCAAGTAGCTGG	AGATTCAAGTGCAGCGCA	CA	667
DB	101888	TTCAAGCATTTCTCTGCTCA	CGCTCCCAAGTAGCTGG	AGATTCAAGTGCAGCGCA	CA	101829
QY	668	CACCCAGCT--TTT	TTT	TTT	TTT	TTT
DB	101828	CACCCAGCTTTT	TTT	TTT	TTT	TTT
QY	726	AGTGCATGATCTAGTTCAGT	TCAGCACTCCGAGGTT	CAAGCAATCTCTG	CC	785
DB	101768	AGTGCATGATCTAGTTCAGT	TCAGCACTCCGAGGTT	CAAGCAATCTCTG	CC	101709

```

786 TCAGCTCCTGAGTGTAGAGATTACAGAAAGGACCTCCAGTTCAGCTAATTTT--G 843
|||||
101708 TCAGCTCCTGAGTGTAGAGATTACAGAAAGGACCTCCAGTTCAGCTAATTTTGG 101649
|||||
844 TATTTTATAGTAGATGCGCTTTTTCGATGTTGGCCATGATGCTGGAACCCGCACT 903
|||||
101648 TGTTTTATAGTAGAGACAGGTTTTCACATGTTAGCAGAGATGGTCTGATCTCTGACCT 101589
|||||
904 CAGGTGATCCGTGCGCTTGGCTCCCAAAGTGTGGGATTTGCAGGCGGTGACCATGCG 963
|||||
101588 CAGGTGATCCGTGCGCTTGGCTCCCAAAGTGTGGGATTTGCAGGCGGTGACCGCG 101529
|||||
964 CCAAGCCTGAGCTA 977
|||||
101528 CCCAGCCACACCCA 101515
|||||
LT 6
Sequence 1, Application US/10071179
Patent No. 6908988
HERED INFORMATION:
APPLICANT: Bousgueleret, Lydie
TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
FIELD OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
FILE REFERENCE: GENSET. 031A
PARENT APPLICATION NUMBER: US/10/071.179
PARENT FILING DATE: 2002-02-07
PARENT APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/345,882
PARENT FILING DATE: EARLIER FILING DATE: 1999-06-30
PARENT APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/091,315
PARENT FILING DATE: EARLIER FILING DATE: 1998-06-30
PARENT APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/111,909
PARENT FILING DATE: EARLIER FILING DATE: 1998-12-10
PARENT OF SEQ ID NOS: 140
SOFTWARE: Patent.psm
Q ID NO 1
LENGTH: 162450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 72794
OTHER INFORMATION: 5-124-273 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 88073
OTHER INFORMATION: 5-127-261 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:

```

NAME/KEY	allele
LOCATION:	103806
OTHER INFORMATION:	5-131-395 : polymorphic base A or T
FEATURE:	
NAME/KEY	allele
LOCATION:	106940
OTHER INFORMATION:	5-133-375 : polymorphic base insertion of A
FEATURE:	
NAME/KEY	allele
LOCATION:	108106
OTHER INFORMATION:	5-135-155 : polymorphic base insertion of A
FEATURE:	
NAME/KEY	allele
LOCATION:	108149
OTHER INFORMATION:	5-135-198 : polymorphic base insertion of GTTT
FEATURE:	
NAME/KEY	allele
LOCATION:	108471
OTHER INFORMATION:	5-136-174 : polymorphic base C or T
FEATURE:	
NAME/KEY	allele
LOCATION:	134134
OTHER INFORMATION:	5-140-120 : polymorphic base C or T
FEATURE:	
NAME/KEY	allele
LOCATION:	134362
OTHER INFORMATION:	5-140-348 : polymorphic base insertion of A
FEATURE:	
NAME/KEY	allele
LOCATION:	134374
OTHER INFORMATION:	5-140-361 : polymorphic base insertion of CA
FEATURE:	
NAME/KEY	allele
LOCATION:	146328
OTHER INFORMATION:	5-143-84 : polymorphic base A or G
FEATURE:	
NAME/KEY	allele
LOCATION:	146345
OTHER INFORMATION:	5-143-101 : polymorphic base A or C
FEATURE:	
NAME/KEY	allele
LOCATION:	160031
OTHER INFORMATION:	5-148-352 : polymorphic base G or T
FEATURE:	
NAME/KEY	allele
LOCATION:	72771..72817
OTHER INFORMATION:	polymorphic fragment 5-124-273 SEQ ID30
FEATURE:	
NAME/KEY	allele
LOCATION:	72771..72817
OTHER INFORMATION:	polymorphic fragment 5-124-273 SEQ ID51
FEATURE:	
NAME/KEY	allele
LOCATION:	88050..88096
OTHER INFORMATION:	polymorphic fragment 5-127-261 SEQ ID31
FEATURE:	
NAME/KEY	allele
LOCATION:	88050..88096
OTHER INFORMATION:	polymorphic fragment 5-127-261 SEQ ID52
FEATURE:	
NAME/KEY	allele
LOCATION:	90819..90865
OTHER INFORMATION:	complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:	
NAME/KEY	allele

```

>OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70/
FEATURE:
>NAME/KEY: allele
>LOCATION: 93690..93736
>OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
>NAME/KEY: allele
>LOCATION: 93690..93736
>OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
>NAME/KEY: allele
>LOCATION: 97099..97145
>OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
>NAME/KEY: allele
>LOCATION: 97099..97145
>OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
>NAME/KEY: allele
>LOCATION: 97130..97177
>OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
>NAME/KEY: allele
>LOCATION: 99075..99121
>OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
>NAME/KEY: allele
>LOCATION: 99075..99121
>OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID35
FEATURE:
>NAME/KEY: allele
>LOCATION: 99094..99140
>OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID36
FEATURE:
>NAME/KEY: allele
>LOCATION: 103783..103828
>OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID37
FEATURE:
>NAME/KEY: allele
>LOCATION: 103783..103828
>OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
>NAME/KEY: allele
>LOCATION: 106918..106966
>OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID38
FEATURE:
>NAME/KEY: allele
>LOCATION: 108084..108130
>OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
>NAME/KEY: allele
>LOCATION: 108127..108177
>OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
>NAME/KEY: allele
>LOCATION: 108127..108177

```

	Query Match	17.3%; Score 323.6; DB 3; Length 162450;	
	Best Local Similarity	81.2%; Pred. No. 1,1e-78;	
	Matches 401; Conservative	0; Mismatches 89; Indels 4; Gaps 2;	
Oy	488 TGAGCTGCTGTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTGAGACAGATCTGCTCT	547	
Dd	102008 TTAGGACACAGATGAATTCTTTTTTTTTTTGGTCCTTTTTTTTGAGACAGAGATCTGCTCT	101949	
Oy	548 GTGCCCCAGGTGTGAGTGCAGTGGCACTGATCTCGCCCATCAGAACTCTGCCCTCCGGA	607	
Dd	101948 GTCCCCAGGTGTGAGTGCAGTGGTCATCTCGGCTCACACTGCACCTTCGCTTCTGGG	101889	
Oy	608 TTCAAAGCATTTCTCTCTCCTCAGCCTCCCAAGTAGTGGATTACAGATGCACGCCACA	667	
Dd	101888 TTCAAAGCATTTCTCTCCTCAGCCTCCCAAGTAGTGGATTACAGATGCACGCCACA	101829	
Oy	668 CACCACAGCT--TTTTTAATTTTGGAGACAGAATCTTGCTCCCTGTACACCAAGCTTGAATAC	725	
Dd	101828 CGCCACACTAATTTTTTTTTTTTTTTTATGATGAGTGTGCTGTGTACACCAAGCTTGAATAC	101769	
Oy	726 AGTGGCANATCTCAGTTCACCTGTGACCTCCACTCCCGGGTTCAAGCAATTCCTCTGCC	785	
Dd	101768 AGTGGCGATCTGCTGCTCACTGMACTCCGCTCCCGGGTTGACCACTTCTCTGCC	101709	
Oy	786 TCAGTCTCTGAGTAGTACGATTACAGAAATGTCACCTCCACGTTTCACTAATTTTT--G	843	
Dd	101708 TCAGTCTCTGAGTAGTACGATTACAGAAATGTCACCTCCACGTTTCACTAATTTTT--G	101649	
Oy	844 TAATTTAGTAGAGATGCGCTTTTGCATGTTGSCCATGCTATGTGGAACCCGGAACCT	903	
Dd	101648 TGATTTAGTAGAGATGCGCTTTTGCATGTTGSCCATGCTATGTGGAATGCTCTGTAACCT	101589	
Oy	904 CAGGTGATCCGCTGCGCTTGCGCTCCCAAGATGCTGGAGATTGACGGGCTGAGCCATGCG	963	
Dd	101588 CAGGTGATCTGCTGCTGCTGCGCTCCCAAGATGCTGGAGATTGACGGGCTGAGCCATGCG	101529	
Oy	964 CCAGGCTGAGCTA 977		
Dd	101528 CCCAGCCACACCCA 101515		
	RESULT 7		
	US-09-949-002-9649/c		
	; Sequence 9649, Application US/09949002		
	; Patent No. 6900016		
	; GENERAL INFORMATION:		
	; APPLICANT: VENTER, J. Craig et al.		
	; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED		
	; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION		
	; FILE REFERENCE: CL000790		
	; CURRENT APPLICATION NUMBER: US/09/949,002		
	; CURRENT FILING DATE: 2000-01-28		
	; PRIOR APPLICATION NUMBER: 60/231,401		
	; PRIOR FILING DATE: 2000-09-08		
	; NUMBER OF SEQ ID NOS: 10823		
	; SOFTWARE: FastSeq for Windows Version 4.0		
	; SEQ ID NO 9649		
	; LENGTH: 601		
	; TYPE: DNA		
	; ORGANISM: Human		
	US-09-949-002-9649		
	Query Match	17.3%; Score 323.2; DB 3; Length 601;	
	Best Local Similarity	82.7%; Pred. No. 8.2e-80;	
	Matches 382; Conservative	1; Mismatches 74; Indels 5; Gaps 1;	
Oy	517 TTTCCTTTTTTTTGTGACACAGATCTTCTCTGTGCTCCAGGCTGAGTGCAGTGCATGA	576	
Dd	511 TTTTCTTTTTTTTGTGAGTGTCTACTCTGTGCTGCTTGAAGGTGAGTGCAGTGCAGAA	452	
Oy	577 TCTCTGCCACTGCACCTGTGCTCCCGGATTCAGAGATTTCTTCTGCTCAGCTCCC	636	

451 TCTCAGGCACTGCAACTCGCTCGCTCCCGGGTTCAAGANTTCTCTGCTCAGCCTCCC 392

637 AAGTAGCTGGGATTACAGGTGTCACGCCACACACCCAGC-----TTTTTTTATTTTGGAG 691

391 AAGTAGCTGGGACTACACAGCAACAGGCCATGCGCCGCGAGCTAATTTTTTTTTTTTTTTTGGAG 332

692 ACAGAGCTTGGCCCTGTGACCCAGAGCTGGAGTAGACAGTGGCATGTATCTCAGTTCACGTGCGA 75

331 ATGAGAGTCTACTCTGTGTCGCTTACGAGGTGAGANTGAGATGTCGCGCAATCTCAGCTCACTGCA 272

752 CTTCCACCTCTCCGGGTTTCAGCAATTCTCTGCTCAAGTCTCTGAGTAGAGTAGAATTAC 811

271 CCTCGCGCTCCCGAGTTCAAGTAGATTCTCTGCTCAGCCCTCGAGTAGCTGGGAATTAC 212

812 AGAAGTAGCACTCCACGCTGAGCTAATTTTTTGTATTTTATAGTAGAGTAGCGCTTTTGCA 872

211 AAGTGGCCACCAACACCGCCGCTAATTTTTTGTATTTTATAGTAGAGTAGCGTCTTGCCA 152

872 TGTTTGACCATGTAGTGTGGAACCCCGGACCTCAGTGTATCCGCTGGCCCTTGGCCTTGCA 931

151 TGTTTGGCCAGGCTGTCTGTGGCACTACTAGCACTCAGTAGATCCACCCGCTGGGCTTGCCA 92

932 AAGGCTGGGATTTGCAAGCGCGTAGCAACCTAGGCGGCGGCGGCGGCTGGA 973

91 AAGTGTCTGGAGTTACAGGCAATGAACCCAGCGGCGGCGGCGGCGGCGGCA 50

```

17 8
9-949-016-13256
quence 13256, Application US/09949016
Ident No. 6812339
HERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13256
LENGTH: 50368
TYPE: DNA
ORGANISM: Human
9-949-016-13256

17.2% Score 321.8; DB 3; Length 50368;
nt Local Similarity 81.8%; Pred.No. 2e-78;
:ches 387; Conservative 0; Mismatches 77; Indels 9; Gaps 1;

507 TTTTTCCTTTCTTTTCTTTTGGAGACAGAGTCTGCTGTGTGCGCCAGGCTGAGTGC 566
|||||
48915 TTGATTTTTTTTTTTCTTTTGGAGAGGGTCTGTGTTGTGTGCCAAGCTGAGTGC 48974
|||||

567 AATGGCAGATCTTGCCCATGCAACTTGCTCCCGAATTCAAGCAATTCTCTGCC 626
|||||
48975 AATGGCAGATCTCGGCTCATGCAACTTGCAATCCCAAGTTCAATGATTTCTCTGCC 49034
|||||

627 TCAGCTCCCAAGTAGCTGGATTACAGTTCACGCCACCAACCAGC-----TT 677
|||||
49035 TTAGCTCTCGTAGTGTGGGATTACAGCAGTAGGACAAAGACACTTGGCTAATTTTTTTT 49094
|||||

678 TTTTATTTTGGAGACAGAGTCTTCCCTGTACACCCAGGCTGAGATCAAGTGGCAATC 737
|||||
49095 TTTTATTTTGGAGACAGAGTCTCTTCTGTACACCCAGGCTGAGATCAATGGCAATC 49154
|||||

```

OY	738	CAATTACACGACACCTCCACCTCCGGGTTCAAGAAATTCCTCTCCTCAGCTCCTGA	797
Db	49155	TCGGTTACACGACCTCCACCTCCAGGTTTAGACAAATTCCTCCTCAGCCTCCCTA	49214
OY	798	GTACCTAGAAATTACAGAAAGTACCTCCACGTTCACTAAATTTTGTATTTTATAGAG	857
Db	49215	GTACCTGGAAATTACAGGTACGACACCACTGTCTGGCTAATTTTGTATTTTATAGAG	49274
OY	858	ATGGGCTTTTGGCATTGTGGCCATGCTAAGTCTGGAAACCCGGACCTCAAGTATCCGCTG	917
Db	49275	ATGGGGTTTGGCATTGTGGCCAGGCTGGTCTAAACTCTGTACCTCAGGTATCCACCT	49334
OY	918	GCTTTGGCCTCCCAAATGTCTGGGAAATTGACAGGCTGAGCCATGCGCCACAGGCC	970
Db	49335	GTCCTGGCCTCTCAAAAGTTCTGGGAAATTACAGGCAATAGCCATCTGGCCCTGGCC	49387

RESULT 9
US-09-94

Sequence 765, A

; Patent No. 6900016
: GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

1. TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL000790

CURRENT APPLICATION NUMBER: US/09/949, 002 ;

CURRENT FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/231,401

PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 10823

; SOFTWARE: FastSeq for Windows Version 4.0.0

```
; SEQ ID NO /65  
; LENGTH: 129554
```

TYPE: DNA ;

ORGANISM: Human

```

; FEATURE:
; NAME/KEY. misc feature

```

LOCATION: (1) .. (129554)

OTHER INFORMATION: n = A, T, C or G


US-09-949-002-765

Query Match	17.1%	Score 320.6	DB 3	Length 129554
-------------	-------	-------------	------	---------------

Best Local Similarity 81.4%; Pred. No. 6.9e-78;

Matches 385; Conservative 0; Mismatches 84; Indels 4; Gaps 1,

515 TTTTCTTTTGTGAGACAGACTCTGCTCTGTCCGCCAGGCTGGAGTGCAGTGGCAT 574



Db 35451 TTTT TTTT TTTT GAGATAGCCTCACTCTGTGCGCCAGGCTGGAGTGCAATGGTGC 3551

575 GATCTCTGCCCACTGCAACCTCTGCGCTCCCGGATTCAAGCGATTCTCCTGCGCTCAGCCTC 634

100

Db 3551 GATCTTGGCTCACTGCAACCTCTGCCCTCTGGGTTCAAGCGATTCTCCTGCGCTCACCTTC 3557

635 CCAAGTACGTCGGATTTACAGGACACCCACACCGAGC-----ATTATTATTATTATTGGA 690

22

Db 35571 CCGAGTTGTGGATTACAGGTGCCACCAACACACCCGGCTGATTTTTCCTTGA 3563

Dur **691 GAGGAGTCTTTGACCAATCAGGAACTAAGTGAGCATTCAGATTACAATGGCG 750**

27

Db 35631 GACGAGTCTTGCTCTGTCGCCCAGGCTGAAGTGCCTGGCGGATCTCTGCTTACTGCA 3569

[illegible]

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

Db 35691 AACTCTGCCCTCCAGGTTCAAGCGATTCTCCTGCTCAGCCTCCTGAGTAGCTGGGATTA 3575

870 **811**

[illegible]

Db 35751 CAGGTGTGTCACCATGCCAGCTAATTTTGTATTTTAGGAGAGACGGGCTTGGC 3581

930 931

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99


```

17 12
9-949-016-14492/c
quence 14492, Application US/09949016
ent No. 6812339
HERAL INFORMATION:
PLICANT: VENTER, J. Craig et al.
FILE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
URRENT APPLICATION NUMBER: US/09/949, 016
URRENT FILING DATE: 2000-04-14
RIOR APPLICATION NUMBER: 60/241,755
RIOR FILING DATE: 2000-10-20
RIOR APPLICATION NUMBER: 60/237,768
RIOR FILING DATE: 2000-10-03
RIOR APPLICATION NUMBER: 60/231,498
RIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
OTMARK: FastSeq for Windows Version 4.0
O ID NO 14492
LENGTH: 7481
TYPE: DNA
ORGANISM: Human
9-949-016-14492

Copy Match      16.9%; Score 315.8; DB 3; Length 7481;
nt Local Similarity 76.5%; Pred. No. 3.4e-77;
ches 400; Conservative 0; Mismatches 122; Indels 1; Gaps 1;

455 TCTGCTTCCTCCCTGAGGCGCAGAGCTGCAGAGATTGAGCTGCTGCTCTTTTTC 514
7437 TCCGCTGCTGCTGCGCTCCCAAAGTCGTGGATTACAGGATAGACACACACCGGCC 7378
455 TTTTCTTTTTCCTTCAGACAGAGCTCTGCTCTGCGCCAGCTGAGTGCAGTGGCAT 574
7377 TTTTCTTTTTCCTTTTGAACAGAGCTCTACCTCTGCCCCAGGCTGAGTGCAGTGGCGC 7318
515 GATCTCTGCGCCACCTGCAACCTCTGCTCTCCCGGATTCAAGGATCTCTGCTCAGCTTC 634
7317 AATCTGCGCTACCTGCACTCTCTGCTCTCTCGGTTCAAGTATCTGTGCTCAGCTTC 7258
635 CCAGTACCTGGGATTACAGGTGACGCGCACACACACCCAGCT..TTTTTATTTTGGAGAC 693
7257 CTGAGTACTGAGACTACAGACGACGACACACAGCTCTGGTAAATTTTTTTTGGAGAC 7198
694 AGAGTCTTGCCCTGTACCCAGCCAGGCTGAGATACAGTGCAGTATGATCTCAGTTCACCTGCGACC 753
7197 GGAAGCTTACTCTGTGGCCAGCTGAGTGCAGATGGATGGATCTTGCTCAGTGCACCC 7138
754 TCCACTCCCGGGTTCAGCAATCTCTGCTCAGCTCTCTAGTATGCTAGATTACAG 813
7137 TCTGCTCCCGAGTTCAAGCATCTCTGCTCAGCTCTCCCAAGTATGCTGGAATTACAG 7078
814 AAGTCACTCCACGCTTCAAGCTAATTTTGAATTTTATAGATAGATGCGCTTTGCGCATG 873
7077 GGGCTGCGACATCAACCCCTGGCTAATTTTGAATTTTATAGTACAGCGGGTTTCATCATTA 7018
874 TTGGCCATGCTAGTGTGAACCCCGGACCTCAGGTATCCGCTGGCTTGAGCTCCCAAA 933
7017 ATGGCAGAGCTGATCTCGAATCTCTGACCTCAGGTGATCTGCGCTTGAGCTCCCAAA 6958
934 GTCGCGGATTTGACGGCGTGAAGCATGCGCGCAGGCGCTGAGCT 976
6957 GTGCTGGATTACAGCGCGTGAAGCAGCGTGTGCGCTTCGCT 6915

```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C0000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 584
; LENGTH: 21308
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-584

Query Match      16.9%; Score 315.8; DB 3; Length 21308;
Best Local Similarity 82.0%; Pred. No. 5.9e-77;
Matches 401; Conservative 0; Mismatches 82; Indels 6; Gaps 3;

QY 505 TTTTTCCTTTCTTTCTTTCTTTTCTTTTGTGAGACAGAGTCTGCTCTGTGCGCCAGGCTGAG 563
DB 3157 TTTCTTTTCTTTTCTTTTCTTTTCTTTTGTGAGACAGAGTTTGGCTCTGTTGCCAGGCTGAG 3098

QY 564 TGCAGTGCATGATCTCTGCCACTGCACACTCTGCTCCCGATTCAAGCGATTCTCT 623
DB 3097 TGCATATGGCATGATCTCGGCTCACACACACTCTGCTCTCAGGTTCAAAAGATTCTCT 3038

QY 624 GCCCAGCCCTCCCAAGTAGCTGGGATTACAGGTGCAAGCCACACACCCAGC---TTTTT 680
DB 3037 GCCCAGCTCTCCGAGTAGCTGGGATTACAGCATGCGGACACAGCCCGGCTTAATTTTT 2978

QY 681 TTATTTTGGAGACAGAGTCTTGGCCCTGTCAACCAGCTGAGTAGACAGTGCATGATCTCA 740
DB 2977 TTTTGTGAGACGGAAGTGTCACTGTCAACCCAGCTGAGTAGTGCATGATCTCA 2918

QY 741 GTTACTGCGCACTTCCACTTCCCGGGTTCAAGCAATTTCTCTGCTCAAGTCTCTTAGTA 800
DB 2917 GCTACTGCAACCTCCACACTCCCGGGTTCAAGTATTTCTCTTACTAGCCTCCGAGTA 2858

QY 801 GCTAGGATTCAGAAAGTGCACCTGCACGTTCAAGCTAATTTTGTATTTTATAGTAGATG 860
DB 2857 GCTGGACTACAGAGTGTGCCACACGCTCAAGCTAATTTTGTATTTTATAGTAGACA 2798

QY 861 CGCTTTTGGCAGATGTGSCCATGCTAGTGTGGAACCCCGAGCTTCAGGTGATCCGCTGACC 920
DB 2797 GGGTTTACCAATATTTGGCCAGGCTGGTCTTGAATCTGTAGCTCA--TGATCTGCCACC 2740

QY 921 TTGGCCCTCCAAAGTGTGGGATTGACAGGGTGAAGCAATGGCCGACGGCTGAGCTACTC 980
DB 2739 TCTGCCCTCCCAAAATGCTGGGATTTGAGAGTGTAGGACCATGTGTGCTCTCTTTT 2680

QY 981 CTTTAGTCT 989
DB 2679 TTTTCTTTT 2671

RESULT 14
US-09-949-002-784/c
; Sequence 784, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: C0000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0

```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocelexeration Ltd.

nucleic - nucleic search, using sw model

on: June 21, 2006, 22:25:04 ; Search time 2431 Seconds

(without alignments)
9452.027 Million cell updates/sec

US-10-502-279-26

1 atctgtctcctagaagctac.....catcccgatcctctag 1870

IDENTITY NUC

Gapop 10.0, Gapext 1.0

number of hits satisfying chosen parameters: 37784340

DB seq length: 0

DB seq length: 200000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

base : Published Applications NA Main:*

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*

2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*

3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq.*

4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*

5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*

6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*

7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*

8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*

9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*

10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*

11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*

12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*

13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*

14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*

15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*

16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	ID	Description
1870	100.0	1870	US-10-502-279-26	Sequence 26, Appl
333.4	17.8	998	US-10-301-480-552165	Sequence 552165,
333.4	17.8	998	US-10-301-480-1165574	Sequence 1165574,
330.8	17.7	33362	US-10-388-838-136	Sequence 36, Appl
326.4	17.5	968	US-10-301-480-605733	Sequence 605733,
326.4	17.5	968	US-10-301-480-1219142	Sequence 1219142,
324.8	17.4	3116	US-10-087-192-1660	Sequence 1660, Ap
324.8	17.4	58822	US-10-052-482-146	Sequence 46, Appl
323.6	17.3	162450	US-10-071-179-1	Sequence 1, Appl
323.6	17.3	162450	US-10-126-704-1	Sequence 1, Appl
323.6	17.3	162450	US-11-132-838-1	Sequence 1, Appl
322.8	17.3	16963	US-10-741-601-5756	Sequence 5756, Ap
322.8	17.3	16963	US-10-741-601-17945	Sequence 17945, A
322.8	17.3	16963	US-10-995-561-13467	Sequence 13467, A
322.4	17.2	296405	US-10-087-192-1036	Sequence 1036, Ap
322.4	17.2	40000	US-10-995-561-13513	Sequence 13513, A
320.6	17.1	146656	US-11-121-086-68	Sequence 68, Appl

C 18	320.4	17.1	4957	7	US-10-074-024-753	Sequence 753, App
C 19	320.4	17.1	4961	7	US-10-074-024-752	Sequence 752, App
C 20	320.4	17.1	41907	7	US-09-967-013-5	Sequence 5, Appl
C 21	320.4	17.1	126001	7	US-10-175-492-13	Sequence 13, Appl
C 22	320.4	17.1	153142	15	US-11-121-086-27	Sequence 27, Appl
C 23	320.4	17.1	201309	16	US-11-114-798-51	Sequence 51, Appl
C 24	320.4	17.1	212022	11	US-10-330-773-102	Sequence 102, App
C 25	320.4	17.1	344805	9	US-10-779-271-1	Sequence 1, Appl
C 26	320.4	17.1	354592	10	US-10-737-082-70	Sequence 70, Appl
C 27	320.4	17.1	354592	10	US-10-765-790-70	Sequence 70, Appl
C 28	320.4	17.1	1080000	10	US-10-928-446A-1	Sequence 1, Appl
C 29	320.4	17.1	1080000	10	US-10-928-446A-181	Sequence 181, App
C 30	320.4	17.1	1080000	10	US-10-928-446A-183	Sequence 183, App
C 31	320.4	17.1	1080000	10	US-10-928-446A-185	Sequence 185, App
C 32	320.4	17.1	1080000	10	US-10-928-446A-187	Sequence 187, App
C 33	320.4	17.1	1080000	10	US-10-928-446A-189	Sequence 189, App
C 34	320.4	17.1	1080000	10	US-10-928-446A-191	Sequence 191, App
C 35	320.4	17.1	1080000	10	US-10-928-446A-193	Sequence 193, App
C 36	320.4	17.1	1080000	10	US-10-928-446A-195	Sequence 195, App
C 37	320.4	17.1	1080000	10	US-10-928-446A-197	Sequence 197, App
C 38	320.4	17.1	1080000	10	US-10-928-446A-199	Sequence 199, App
C 39	320.4	17.1	1080000	10	US-10-928-446A-201	Sequence 201, App
C 40	320.2	17.1	216929	8	US-10-741-601-5727	Sequence 5727, Ap
C 41	319.8	17.1	6093	7	US-10-074-024-535	Sequence 535, App
C 42	318	17.0	9649	3	US-09-956-712-10	Sequence 10, Appl
C 43	318	17.0	9649	8	US-10-633-913-10	Sequence 10, Appl
C 44	317.4	17.0	120096	15	US-11-121-086-24	Sequence 24, Appl
C 45	316	16.9	34875	9	US-10-775-169-316	Sequence 316, App

ALIGNMENTS

US-10-502-279-26

Sequence 26, Appl

Publication No. US20050084840A1

GENERAL INFORMATION:

APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.

APPLICANT: Hideki ENDOH

APPLICANT: Ryouke NAKANO

APPLICANT: Eiji KUROSAKI

APPLICANT: Mayuki KATO

APPLICANT: Hiroyuki YOKOTA

APPLICANT: Kazumori INABE

TITLE OF INVENTION: METHOD FOR SCREENING A DRUG AMELIORATING INSULIN RESISTANCE

FILE REFERENCE: 082704

CURRENT APPLICATION NUMBER: US/10/502,279

CURRENT FILING DATE: 2004-07-23

PRIOR APPLICATION NUMBER: JP 2002-013721

PRIOR FILING DATE: 2002-01-23

PRIOR APPLICATION NUMBER: JP 2002-257703

PRIOR FILING DATE: 2002-09-03

NUMBER OF SEQ ID NOS: 28

SOFTWARE: PatentIn version 3.1

SEQ ID NO 26

LENGTH: 1870

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: promoter

LOCATION: (1)..(1870)

OTHER INFORMATION:

US-10-502-279-26

Query Match 100.0%; Score 1870; DB 10; Length 1870;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCTGTCTCCTAAGATACCAAGCAATCTACAGGCTCTGAAATACGCTTCT 60

1 ATCTGTCTCCTAAGATACCAAGCAATCTACAGGCTCTGAAATACGCTTCT 60

61 AGTTCTAGATTTCACTCTCATGTCTGTGGTATGCTCCCTTATTTGCCAGCTCCATG 120
61 AGTTCTAGATTTCACTCTCATGTCTGTGGTATGCTCCCTTATTTGCCAGCTCCATG 120
121 GCAAGTTGGCTCTCTCTGAGCTCTTGGATGAACCTGATTTTCANATGCTCATAGGCGCA 180
121 GCAAGTTGGCTCTCTCTGAGCTCTTGGATGAACCTGATTTTCANATGCTCATAGGCGCA 180
121 GCAAGTTGGCTCTCTCTGAGCTCTTGGATGAACCTGATTTTCANATGCTCATAGGCGCA 180
181 GGATATGAGCTAGAGAAAGTTACATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
181 GGATATGAGCTAGAGAAAGTTACATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
241 CATTTCTGCTATTTCTTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
241 CATTTCTGCTATTTCTTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
301 AAAAATTTAGAGTCAAGAGTATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
301 AAAAATTTAGAGTCAAGAGTATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
361 CTGGGAAAAATTTCAAG 420
361 CTGGGAAAAATTTCAAG 420
421 GGGTCTCAATCAAGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
421 GGGTCTCAATCAAGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
481 CAGAGATGAGAGTGTGCTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 540
481 CAGAGATGAGAGTGTGCTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 540
541 TTGCTCTGTGCGCCAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
541 TTGCTCTGTGCGCCAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
541 TTGCTCTGTGCGCCAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
601 TCCCGGATTTCAAGCGATTTCTCTGCTCAGCTCCAGCTCCAGAGAGAGAGAGAGAGAGAG 660
601 TCCCGGATTTCAAGCGATTTCTCTGCTCAGCTCCAGCTCCAGAGAGAGAGAGAGAGAGAG 660
661 GGCACACACCCAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 720
661 GGCACACACCCAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 720
721 AGTACAGTGGAGATCTAGTTCACTGCGACCTCCAGCTCCGAGGTTCAAGCAATTTCTC 780
721 AGTACAGTGGAGATCTAGTTCACTGCGACCTCCAGCTCCGAGGTTCAAGCAATTTCTC 780
781 CTGCTCAAGTCTCTGAG 840
781 CTGCTCAAGTCTCTGAG 840
841 TTGATATTTTATGAGAGATGCGCTTTTGGCATTTTGGCATTTTGGCATTTTGGCATTTTGG 900
841 TTGATATTTTATGAGAGATGCGCTTTTGGCATTTTGGCATTTTGGCATTTTGGCATTTTGG 900
901 CCTCAGAGTATTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
901 CCTCAGAGTATTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
961 GCGCCAGGCTGAGCTACTCTTTTATGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
961 GCGCCAGGCTGAGCTACTCTTTTATGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
1021 CTTTACATGCAATCTCTCCCTAGTCCCAAGGCTCTCTGAGAGAGAGAGAGAGAGAGAGAG 1080
1021 CTTTACATGCAATCTCTCCCTAGTCCCAAGGCTCTCTGAGAGAGAGAGAGAGAGAGAGAG 1080
1081 CATCTCTCAATCAG 1140
1081 CATCTCTCAATCAG 1140
1141 TAACTCTTCCCGAGAACTCAGTAGGCTCTGATGAGCATCAGTAAGGGCAATTTTCAATA 1200

Db 1141 TAACTCTTCCCGAGAACTCAGTAGGCTCTGATGAGATCAGTAAGGGCAATTTTCAATA 1200
Qy 1201 AGCAACAGAGGTTCTGCTGCTTTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 AGCAACAGAGGTTCTGCTGCTTTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Qy 1261 TGAGGAG 1320
Db 1261 TGAGGAG 1320
Qy 1321 GGGCCCCCAAGAGTCTGAAATCTAGTGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 GGGCCCCCAAGAGTCTGAAATCTAGTGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 TCTGTTTATGAGCTAGAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 TCTGTTTATGAGCTAGAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy 1441 CCCACCTTAACTCCATTTCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 CCCACCTTAACTCCATTTCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Qy 1501 TGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
Db 1501 TGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
Qy 1561 CTACCCCATCAAGGAG 1620
Db 1561 CTACCCCATCAAGGAG 1620
Qy 1621 CCTTCCCGCTCATATTCCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 CCTTCCCGCTCATATTCCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Qy 1681 CAAAGGAG 1740
Db 1681 CAAAGGAG 1740
Qy 1741 ACACCCCGGCTGAGTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 1741 ACACCCCGGCTGAGTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Qy 1801 GCGGGGCGCTTACCATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db 1801 GCGGGGCGCTTACCATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Qy 1861 TCCCTGCTAG 1870
Db 1861 TCCCTGCTAG 1870

RESULT 2
US-10-301-480-552165/c
; Sequence 552165, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1228618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 552165
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Homo sapien


```

757 ACCCTCCGGGTTCAAGCAATTCCTGCTCAGTCTCTGAGTAGTAGATTAACAAG 816
11946 GCCGCCGGGATTCAGCGCATTCCTGCTCAGCCCTCCGAGTAGTAGATTAACAAG 11887
817 TCCACTCCAGCTTCAAGCAATTCCTGCTCAGTCTCTGAGTAGTAGATTAACAAG 876
11886 TGTCCACCATGCGCGGCTTAATTTTGTATTTTGTAGTAGTAGATTAACAAG 11827
877 GCCATGCTAGTCTGGAACCCCGGACCTCAGTGTATCCGCTGCTGCTCCAAAGT 936
11826 GCAAGGCTGTCTTGAATCTCTGAGTATCCCTCAGGATATCCCTCAGGCTCCAAAGT 11767
937 CTGGATTGCAAGGCTGAGCCATCGCCCAAGCC 970
11766 CTGGATTGCAAGGCTGAGCCATCGCCCAAGCC 11733

JF 5
1-301-480-605733
Sequence 605733, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE OF INVENTION: In the Human Genome
REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
) ID NO 605733
LENGTH: 968
TYPE: DNA
ORGANISM: Homo sapien
1-301-480-605733

Query Match      17.5%; Score 326.4; DB 12; Length 968;
Local Similarity 83.5%; Pred. No. 2.8e-87;
Matches 396; Conservative 0; Mismatches 71; Indels 7; Gaps 2;

503 TTTTCTTTCTTTCTTTCTTTCTTTTGAAGACAGAGTCTGCTGCTGCTGCTGGA 562
473 TTGATTTTCTTTCTTTCTTTCTTTTGAAGACAGAGTCTGCTGCTGCTGCTGGA 532
563 GTGAGTGGCATGATCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGA 622
533 GTACAGTGGCATGATCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGA 592
623 TGCCTCAGCTCCCAAGTACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 677
593 TGCCTCAGCTCCCAAGTACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 652
678 TTTTATTTTGAAGACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
653 TTTTATTTTGAAGACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 712
738 TCAGTTCATGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 797
713 TCGGTTCAAGTCAAGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 772
798 GTAGTAGGATTACAGAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 857
773 GTAGTAGGATTACAGAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 832
858 ATGCGCTTTTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGG 917
833 ATGCGCTTTTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGG 890
918 GCCTTGCTCCCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCT 971

```

```

Db      891 GCCTTGCTCCCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCT 944

RESULT 6
US-10-301-480-1219142
Sequence 1219142, Application US/10301480
Publication No. US20060057564A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
FILE OF INVENTION: In the Human Genome
REFERENCE: 108827.137
CURRENT APPLICATION NUMBER: US/10/301,480
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/215,598
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,695
PRIOR FILING DATE: 2001-08-10
NUMBER OF SEQ ID NOS: 1226818
SOFTWARE: FastSeq for Windows Version 4.0
) SEQ ID NO 1219142
LENGTH: 968
TYPE: DNA
ORGANISM: Homo sapien
US-10-301-480-1219142

Query Match      17.5%; Score 326.4; DB 12; Length 968;
Best Local Similarity 83.5%; Pred. No. 2.8e-87;
Matches 396; Conservative 0; Mismatches 71; Indels 7; Gaps 2;

QY 503 TTTTCTTTCTTTCTTTCTTTCTTTTGAAGACAGAGTCTGCTGCTGCTGCTGGA 562
DB 473 TTGATTTTCTTTCTTTCTTTCTTTTGAAGACAGAGTCTGCTGCTGCTGCTGGA 532
QY 563 GTGAGTGGCATGATCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGA 622
DB 533 GTACAGTGGCATGATCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGA 592
QY 623 TGCCTCAGCTCCCAAGTACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 677
DB 593 TGCCTCAGCTCCCAAGTACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 652
QY 678 TTTTATTTTGAAGACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
DB 653 TTTTATTTTGAAGACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 712
QY 738 TCAGTTCATGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 797
DB 713 TCGGTTCAAGTCAAGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 772
QY 798 GTAGTAGGATTACAGAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 857
DB 773 GTAGTAGGATTACAGAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 832
QY 858 ATGCGCTTTTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGG 917
DB 833 ATGCGCTTTTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGG 890
QY 918 GCCTTGCTCCCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCT 971
DB 891 GCCTTGCTCCCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCAAGTGTGGCT 944

RESULT 7
US-10-087-192-1660
Sequence 1660, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
CANCER

```


ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: allele
 LOCATION: 72794
 OTHER INFORMATION: 5-124-273 : polymorphic base A or G
 FEATURE:
 NAME/KEY: allele
 LOCATION: 88073
 OTHER INFORMATION: 5-127-261 : polymorphic base A or C
 FEATURE:
 NAME/KEY: allele
 LOCATION: 90842
 OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
 FEATURE:
 NAME/KEY: allele
 LOCATION: 93714
 OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
 FEATURE:
 NAME/KEY: allele
 LOCATION: 97122
 OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
 FEATURE:
 NAME/KEY: allele
 LOCATION: 97152
 OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
 FEATURE:
 NAME/KEY: allele
 LOCATION: 99098
 OTHER INFORMATION: 5-130-257 : polymorphic base A or G
 FEATURE:
 NAME/KEY: allele
 LOCATION: 99117
 OTHER INFORMATION: 5-130-276 : polymorphic base A or G
 FEATURE:
 NAME/KEY: allele
 LOCATION: 103806
 OTHER INFORMATION: 5-131-395 : polymorphic base A or T
 FEATURE:
 NAME/KEY: allele
 LOCATION: 106940
 OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
 FEATURE:
 NAME/KEY: allele
 LOCATION: 108106
 OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
 FEATURE:
 NAME/KEY: allele
 LOCATION: 108149
 OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
 FEATURE:
 NAME/KEY: allele
 LOCATION: 108308
 OTHER INFORMATION: 5-135-357 : polymorphic base A or G
 FEATURE:
 NAME/KEY: allele
 LOCATION: 108471
 OTHER INFORMATION: 5-136-174 : polymorphic base C or T
 FEATURE:
 NAME/KEY: allele
 LOCATION: 134134
 OTHER INFORMATION: 5-140-120 : polymorphic base C or T
 FEATURE:
 NAME/KEY: allele
 LOCATION: 134362
 OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
 FEATURE:
 NAME/KEY: allele
 LOCATION: 134374
 OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
 FEATURE:
 NAME/KEY: allele
 LOCATION: 146328
 OTHER INFORMATION: 5-143-84 : polymorphic base A or G

FEATURE:
 NAME/KEY: allele
 LOCATION: 146345
 OTHER INFORMATION: 5-143-101 : polymorphic base A or C
 FEATURE:
 NAME/KEY: allele
 LOCATION: 150329
 OTHER INFORMATION: 5-145-24 : polymorphic base A or G
 FEATURE:
 NAME/KEY: allele
 LOCATION: 160031
 OTHER INFORMATION: 5-148-352 : polymorphic base G or T
 FEATURE:
 NAME/KEY: allele
 LOCATION: 72771..72817
 OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
 FEATURE:
 NAME/KEY: allele
 LOCATION: 72771..72817
 OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID31
 FEATURE:
 NAME/KEY: allele
 LOCATION: 88050..88096
 OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID32
 FEATURE:
 NAME/KEY: allele
 LOCATION: 88050..88096
 OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID33
 FEATURE:
 NAME/KEY: allele
 LOCATION: 90819..90865
 OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
 FEATURE:
 NAME/KEY: allele
 LOCATION: 90819..90865
 OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID50
 FEATURE:
 NAME/KEY: allele
 LOCATION: 93690..93736
 OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID51
 FEATURE:
 NAME/KEY: allele
 LOCATION: 93690..93736
 OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID52
 FEATURE:
 NAME/KEY: allele
 LOCATION: 97099..97145
 OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
 FEATURE:
 NAME/KEY: allele
 LOCATION: 97099..97145
 OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID72
 FEATURE:
 NAME/KEY: allele
 LOCATION: 97130..97177
 OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
 FEATURE:
 NAME/KEY: allele
 LOCATION: 97130..97177
 OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID34
 FEATURE:
 NAME/KEY: allele
 LOCATION: 99075..99121
 OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
 FEATURE:
 NAME/KEY: allele
 LOCATION: 99075..99121
 OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID35
 FEATURE:
 NAME/KEY: allele
 LOCATION: 99094..99140
 OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
 FEATURE:

```
NAME/KEY: allele  
LOCATION: .99094..99140  
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56  
FEATURE:  
NAME/KEY: allele  
LOCATION: .103783..103828  
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57  
FEATURE:  
NAME/KEY: allele  
LOCATION: .103783..103828  
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57  
FEATURE:  
NAME/KEY: allele  
LOCATION: .106918..106966  
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID57  
FEATURE:  
NAME/KEY: allele  
LOCATION: .106918..106966  
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58  
FEATURE:  
NAME/KEY: allele  
LOCATION: .106918..106966  
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58  
FEATURE:  
NAME/KEY: allele  
LOCATION: .108084..108130  
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59  
FEATURE:  
NAME/KEY: allele  
LOCATION: .108127..108177  
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID59  
FEATURE:  
NAME/KEY: allele  
LOCATION: .108127..108177
```

key Match 17.3%; Score 323.6; DB 6; Length 162450;
 Local Similarity 81.2%; Pred No. 2.8e-85;
 nches 401; Conservative 0; Mismatches 89; Indels 4; Gaps 2;

```
488 TGAGCTGTGTCGTCTTTTATTTTTTCTTTTCTTTTATTGAAGAAGTCGTGCTC 547  
120208 TAAAGAGCAAGTAGTAATCTTTTGGTGTCGTTTTGAACAAGTCGTGCTC 101949  
  
548 GTCCGCCAAGCGTGAAGTGCATCATCTGCCCCACCTGCAACTCTGCTCCCGGA 607  
101948 GTCCGCCAAGCGTGAAGTGCATCATCTGCCCCACCTGCAACTCTGCTCCG 101889  
  
608 TTCAAGCCAATTCCTCGCTCACGCTCCCAAGTAGCTGGGATTACAGTACAGSCCA 667  
101888 TTCAGCAATCTCGCTGCTCACGCTCCCTAGTAGCTGGGATTACAGSCCA 101829  
  
668 CACCACAGCT--TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCAACCAGGCTGAGTAC 725  
101828 CGCCACAGCTAATTTTTTTTTTTTGAAGAGTGTGCTGTGTACCAAGGCTGAGTGC 101769  
  
726 AGTGCAAGTATCTCACTTCAGCTGCACCTCCCGGTTCAAGCAATTCCTGTC 785  
101768 AGTGCAAGTATCTCGCTCAGCTGCAGAGCTCCGCTCCCGGGTTGACACATTCCTGTC 101709  
  
786 TCAGTCTCTGAGTAGCTAAGATTACAAAAGTGACCTCCACGTTCAAGCTAATTTT -G 843  
101708 TCAGCTCTCTGAGTAGCTGAGACTAACAGAGCTGCACCAACATGCGGCTAATTTTTTGG 101649  
  
844 TATTTTATAGTAGAGTGGCTTTTGGCAATGTTGGCAATGTAAGCTGGAACCCCGAGCT 903  
101648 TATTTTATAGTAGAGCAAGGTTTACCAATGTTAGCAAGATGATCTGATCTCTGACT 101589  
  
904 CAGGTATCCGCTGCGCTTGCGCTCCCAAAGTGTCTGAGATTGCAAGSGTGAAGCATCGCG 963  
101588 CAGGTATCTGCTCTCTTGCGCTCCCAAAGTGTCTGAGATTACAGGCGTGACCGCG 101529  
  
964 CCAGGCTGAGCTA 977
```

```

Db      101528  CCCAGCCACACCCA 101515

RESULT 10
US-10-126-704-1/c
; Sequence 1, Application US/10126704
; Publication No. US20030170647A1
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBB-7)
; FILE REFERENCE: 44.US5.DIV
; CURRENT FILING DATE: 2002-04-20
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93714
; OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97122
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97152
; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99098
; OTHER INFORMATION: 5-130-257 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99117
; OTHER INFORMATION: 5-130-276 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103806
; OTHER INFORMATION: 5-131-395 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106940
; OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108106
; OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108149
; OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
; FEATURE:

```


101888 TTCAAGAAATCGCTGCTCAGCTCTCTAGTAGTGGAGATTAACAGCACTGCTCCACCA 101829
668 CACCCAGCT--TTTTTATTTTGGAGACAGAGTCTTCCCTGTCACCCAGGCTGGAGTAC 725
101828 CGCCAGCTAAATTTTTTTTTTTTTTAAATGAGTCTGCTCTGTCACCCAGGCTGGAGTGC 101769
726 AGTGCATGATCTCAGTTCACCTGCACTCCACCTCCCGGGATTCAAGCAATTCCTCTGCC 785
101768 AGTGGGTGATCTCGCTGCTCACTGCAAGCTCCGCTCCCGGGTTAGACCAATTCCTCTGCC 101709
786 TCAGTCTCTGAGTAGTCTAGATTAACAGAGTCACTCCACCTGCTCAATTTT--G 843
101708 TCAGCTCTCTGAGTAGTCTAGATTAACAGAGTCACTCCACCTGCTCAATTTT--G 101649
844 TATTTTATGAGATGATCGCTTTTGGCATGTTGGCGCATGCTAGTCTGAAACCCCGGAGCT 903
101648 TATTTTATGAGATGATCGCTTTTGGCATGTTGGCGCATGCTAGTCTGAAACCCCGGAGCT 101589
904 CAGGTGATCCGCTGCTGCTTGGCTCCCAAGTCTGGGATTTGCAAGGCGTGAAGCCATGCG 963
101588 CAGGTGATCTGCTGCTGCTTGGCTCCCAAGTCTGGGATTTGCAAGGCGTGAAGCCATGCG 101529
964 CCAGGCTGAGCTA 977
101528 CCCAGCCACACCA 101515

IT 11
--132-838-1/C
quence 1, Application US/11132838
blication No. US20050221371A1
HERAL INFORMATION:
PLICANT: Bougueleret, Lydie
TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING
SITE OF INVENTION: PROTEIN (RBP-7)
TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
LE REFERENCE: GENSET.031A
URRENT APPLICATION NUMBER: US/11/132,838
URRENT FILING DATE: 2005-05-19
RIOR APPLICATION NUMBER: US/09/345,882
RIOR FILING DATE: 1999-06-30
RIOR APPLICATION NUMBER: US 60/091,315
RIOR FILING DATE: 1998-06-30
RIOR APPLICATION NUMBER: US 60/111,909
RIOR FILING DATE: 1998-12-10
UMBER OF SEQ ID NOS: 140
FTWARE: Patent.pm
Q ID NO 1
LENGTH: 162450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 72794
OTHER INFORMATION: 5-124-273 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 88073
OTHER INFORMATION: 5-127-261 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:

NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele

```

:LOCATION: 88050..88096
:OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
:NAME/KEY: allele
:LOCATION: 88050..88096
:OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
:NAME/KEY: allele
:LOCATION: 90819..90865
:OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID494
FEATURE:
:NAME/KEY: allele
:LOCATION: 90819..90865
:OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID700
FEATURE:
:NAME/KEY: allele
:LOCATION: 90819..90865
:OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID700
FEATURE:
:NAME/KEY: allele
:LOCATION: 93690..93736
:OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
:NAME/KEY: allele
:LOCATION: 93690..93736
:OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
:NAME/KEY: allele
:LOCATION: 97099..97145
:OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
:NAME/KEY: allele
:LOCATION: 97099..97145
:OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
:NAME/KEY: allele
:LOCATION: 97130..97177
:OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
:NAME/KEY: allele
:LOCATION: 97130..97177
:OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
:NAME/KEY: allele
:LOCATION: 99075..99121
:OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
:NAME/KEY: allele
:LOCATION: 99075..99121
:OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
:NAME/KEY: allele
:LOCATION: 99094..99140
:OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
:NAME/KEY: allele
:LOCATION: 99094..99140
:OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
:NAME/KEY: allele
:LOCATION: 103783..103828
:OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
:NAME/KEY: allele
:LOCATION: 103783..103828
:OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
:NAME/KEY: allele
:LOCATION: 106918..106966
:OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
:NAME/KEY: allele
:LOCATION: 106918..106966
:OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
:NAME/KEY: allele
:LOCATION: 108084..108130

```

```

OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID33
FEATURE:
NAME/KEY: allele

```

Query Match	17.3%;	Score 323.6;	DB 13;	Length 162450;
Best Local Similarity	81.2%;	Pred. No. 2.8e-85;		
Matches 401;	Conservative 0;	Mismatches 89;	Indels 4;	Gaps 2

[illegible]

```

RESULT 12
US-10-741-601-5756/c
: Sequence 5756, Application US/10741601
: Publication No. US20040166519A1
: GENERAL INFORMATION:
: APPLICANT: CARCILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: STENOISIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CE001500
: CURRENT APPLICATION NUMBER: US/10/741,601
: CURRENT FILING DATE: 2003-12-22
: NUMBER OF SEQ ID NOS: 26415
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 5756
: LENGTH: 16363
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-741-601-5756

```

Query Match 17.3%; Score 322.8; DB 8; Length 16963;

```

"Local Similarity 82.8%; Pred. No. 1.5e-85;
Indels 11; Gaps 2;
Matches 394; Conservative 2; Mismatches 69; Indels 11; Gaps 2;

506 TTTTTCCTTTCTTTTCTTTTGAAGAGAGTCTGTCTGTGCGCCAGGCTGAGTG 565
|||||
767 TTTGTTTTTTTTTTTTTTTTTGGAGATGAGTCTGTCTGTGAGCCAGCTGAGTG 7708
|||||
566 CAGTGGCATGATCTGTGCGCACTGCAACCTGTGCTTCCCGATTCAGCGATTCCTGTC 625
|||||
7707 CAGTGGCATGATCTGTGCGCACTGCAACCACTTCCCGGTTCAAGGATTCCTGTC 7648
|||||
626 CTCAGCTCCCAAGTAGCTGGATTAAGGTGACGCGACCAACCAAGCTTTTATT 685
|||||
7647 CTCAGCTCCCAAGTAGCTGGATTAAGGTGACCAACCAAGCTTTTATT 7588
|||||
686 TTG-----GAGACAGAGTCTGTGCGCTGTCAACCGAGCTGAGTGAATGAT 726
|||||
7587 TTGTTTGTGTTGAGAGCGAGTCTGTGCTGTGTGACCGAGCTGAGTGAATGAT 7528
|||||
737 CTCAGTTCACCTGAGCTCCACCTCCGCGGTTCAAGCAATTCCTGCTCAGTCTCTG 796
|||||
7527 CTCGGCTCACTGCACTCCGCTCCGCGGTTCAAGTATTCCTGCTCAGCTCCG 7468
|||||
797 AGTAGCTAGGATTAAGAGATGCACTCCAGCTTCAGCTAA--TTTTGATTTTAA 854
|||||
7467 AGTAGCTGGGCTACAGGCGCGCTGCTCCACCTGCTAAATTTTGTATTTTAA 7408
|||||
855 GAGATGCGCTTTTGCATGTTGCGCATGAGTGTGAGACCGCGAGCTCAGGTATCCG 914
|||||
7407 GAGACGAAGTTTACCATGTTGCGCAGGCTGTTTCACTTCCAGCTCAAGTATGAT 7348
|||||
915 CTGCGCTTGCGCTCCCAAGAGTGTGAGATTCAGGCTGAGCACTGCGCCAGGCC 970
|||||
7347 CCGGCTCGGCTCCCAAGTGTGAGATTCAGGCTGAGCACTGCGCCAGGCC 7292
|||||

1.17.13
0-741-600-17945/c
Sequence 17945, Application US/10741600
Publication No. US20050026169A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
DISEASE OF INHERITION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001499
CURRENT FILING DATE: 2003-12-22
NUMBER OF SEQ ID NOS: 7397
SOFTWARE: FASTSEQ for Windows Version 4.0
Q ID NO 17945
LENGTH: 16963
TYPE: DNA
ORGANISM: Homo sapiens
0-741-600-17945

Query Match 17.3%; Score 322.8; DB 9; Length 16963;
"Local Similarity 82.8%; Pred. No. 1.5e-85;
Indels 11; Gaps 2;
Matches 394; Conservative 2; Mismatches 69; Indels 11; Gaps 2;

506 TTTTTCCTTTCTTTTCTTTTGAAGAGAGTCTGTCTGTGCGCCAGGCTGAGTG 565
|||||
767 TTTGTTTTTTTTTTTTTTTTTGGAGATGAGTCTGTCTGTGAGCCAGCTGAGTG 7708
|||||
566 CAGTGGCATGATCTGTGCGCACTGCAACCTGTGCTTCCCGATTCAGCGATTCCTGTC 625
|||||
7707 CAGTGGCATGATCTGTGCGCACTGCAACCACTTCCCGGTTCAAGGATTCCTGTC 7648
|||||
626 CTCAGCTCCCAAGTAGCTGGATTAAGGTGACGCGACCAACCAAGCTTTTATT 685
|||||
7647 CTCAGCTCCCAAGTAGCTGGATTAAGGTGACCAACCAAGCTTTTATT 7588
|||||
686 TTG-----GAGACAGAGTCTGTGCGCTGTCAACCGAGCTGAGTGAATGAT 726
|||||
7587 TTGTTTGTGTTGAGAGCGAGTCTGTGCTGTGTGACCGAGCTGAGTGAATGAT 7528
|||||
737 CTCAGTTCACCTGAGCTCCACCTCCGCGGTTCAAGCAATTCCTGCTCAGTCTCTG 796
|||||
7527 CTCGGCTCACTGCACTCCGCTCCGCGGTTCAAGTATTCCTGCTCAGCTCCG 7468
|||||
797 AGTAGCTAGGATTAAGAGATGCACTCCAGCTTCAGCTAA--TTTTGATTTTAA 854
|||||
7467 AGTAGCTGGGCTACAGGCGCGCTGCTCCACCTGCTAAATTTTGTATTTTAA 7408
|||||
855 GAGATGCGCTTTTGCATGTTGCGCATGAGTGTGAGACCGCGAGCTCAGGTATCCG 914
|||||
7407 GAGACGAAGTTTACCATGTTGCGCAGGCTGTTTCACTTCCAGCTCAAGTATGAT 7348
|||||
915 CTGCGCTTGCGCTCCCAAGAGTGTGAGATTCAGGCTGAGCACTGCGCCAGGCC 970
|||||
7347 CCGGCTCGGCTCCCAAGTGTGAGATTCAGGCTGAGCACTGCGCCAGGCC 7292
|||||

```

```

Db 7587 TTGTTTGTGTTGAGAGCGAGTCTGTGCTGTGTGACCGAGCTGAGTGAATGAT 7528
|||||
Qy 737 CTCAGTTCACCTGAGCTCCACCTCCGCGGTTCAAGCAATTCCTGCTCAGTCTCTG 796
|||||
Db 7527 CTCGGCTCACTGCACTCCGCTCCGCGGTTCAAGTATTCCTGCTCAGCTCCG 7468
|||||
Qy 797 AGTAGCTAGGATTAAGAGATGCACTCCAGCTTCAGCTAA--TTTTGATTTTAA 854
|||||
Db 7467 AGTAGCTGGGCTACAGGCGCGCTGCTCCACCTGCTAAATTTTGTATTTTAA 7408
|||||
Qy 855 GAGATGCGCTTTTGCATGTTGCGCATGAGTGTGAGACCGCGAGCTCAGGTATCCG 914
|||||
Db 7407 GAGACGAAGTTTACCATGTTGCGCAGGCTGTTTCACTTCCAGCTCAAGTATGAT 7348
|||||
Qy 915 CTGCGCTTGCGCTCCCAAGAGTGTGAGATTCAGGCTGAGCACTGCGCCAGGCC 970
|||||
Db 7347 CCGGCTCGGCTCCCAAGTGTGAGATTCAGGCTGAGCACTGCGCCAGGCC 7292
|||||

RESULT 14
US-10-995-561-13467/c
Sequence 13467, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
DISEASE OF INHERITION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: CL001559
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FASTSEQ for Windows Version 4.0
Q ID NO 13467
LENGTH: 16963
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-13467

Query Match 17.3%; Score 322.8; DB 10; Length 16963;
"Local Similarity 82.8%; Pred. No. 1.5e-85;
Indels 11; Gaps 2;
Matches 394; Conservative 2; Mismatches 69; Indels 11; Gaps 2;

506 TTTTTCCTTTCTTTTCTTTTGAAGAGAGTCTGTCTGTGCGCCAGGCTGAGTG 565
|||||
767 TTTGTTTTTTTTTTTTTTTTTGGAGATGAGTCTGTCTGTGAGCCAGCTGAGTG 7708
|||||
566 CAGTGGCATGATCTGTGCGCACTGCAACCTGTGCTTCCCGATTCAGCGATTCCTGTC 625
|||||
7707 CAGTGGCATGATCTGTGCGCACTGCAACCACTTCCCGGTTCAAGGATTCCTGTC 7648
|||||
626 CTCAGCTCCCAAGTAGCTGGATTAAGGTGACGCGACCAACCAAGCTTTTATT 685
|||||
7647 CTCAGCTCCCAAGTAGCTGGATTAAGGTGACCAACCAAGCTTTTATT 7588
|||||
686 TTG-----GAGACAGAGTCTGTGCGCTGTCAACCGAGCTGAGTGAATGAT 726
|||||
7587 TTGTTTGTGTTGAGAGCGAGTCTGTGCTGTGTGACCGAGCTGAGTGAATGAT 7528
|||||
737 CTCAGTTCACCTGAGCTCCACCTCCGCGGTTCAAGCAATTCCTGCTCAGTCTCTG 796
|||||
7527 CTCGGCTCACTGCACTCCGCTCCGCGGTTCAAGTATTCCTGCTCAGCTCCG 7468
|||||
797 AGTAGCTAGGATTAAGAGATGCACTCCAGCTTCAGCTAA--TTTTGATTTTAA 854
|||||
7467 AGTAGCTGGGCTACAGGCGCGCTGCTCCACCTGCTAAATTTTGTATTTTAA 7408
|||||
855 GAGATGCGCTTTTGCATGTTGCGCATGAGTGTGAGACCGCGAGCTCAGGTATCCG 914
|||||
7407 GAGACGAAGTTTACCATGTTGCGCAGGCTGTTTCACTTCCAGCTCAAGTATGAT 7348
|||||
915 CTGCGCTTGCGCTCCCAAGAGTGTGAGATTCAGGCTGAGCACTGCGCCAGGCC 970
|||||
7347 CCGGCTCGGCTCCCAAGTGTGAGATTCAGGCTGAGCACTGCGCCAGGCC 7292
|||||

```

7347 CCGGCTGGGCTCCCAAGTGTGAGTTACAGGCTGAGCCACCTGCGCGCC 7292

Job time : 2441 secs

15

1-087-192-1036
 1036, Application US/10087192
 1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

1036, Application No. US20020182586A1

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

nucleic - nucleic search, using sw model

on: June 21, 2006, 22:25:40 ; Search time 59 Seconds
(without alignments)
7598.058 Million cell updates/sec

US-10-502-279-26

het score: 1870

ance: 1 atctgtctctagaagac.....catcccgagctctctag 1870

ing table: IDENTITY NUC

Gapop 10_0, Gapext 1.0

ched: 296510 seqs, 119862409 residues

number of hits satisfying chosen parameters: 593020

num DB seq length: 0
num DB seq length: 200000000

Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

base : Published Applications NA.New.*
1: /EMC_Celerra_SIDS3/pcodata/1/pubpna/US09_NEW_PUB.seq.*
2: /EMC_Celerra_SIDS3/pcodata/1/pubpna/US06_NEW_PUB.seq.*
3: /EMC_Celerra_SIDS3/pcodata/1/pubpna/US07_NEW_PUB.seq.*
4: /EMC_Celerra_SIDS3/pcodata/1/pubpna/US08_NEW_PUB.seq.*
5: /EMC_Celerra_SIDS3/pcodata/1/pubpna/PCT_NEW_PUB.seq.*
6: /EMC_Celerra_SIDS3/pcodata/1/pubpna/US10_NEW_PUB.seq.*
7: /EMC_Celerra_SIDS3/pcodata/1/pubpna/US11_NEW_PUB.seq.*
8: /EMC_Celerra_SIDS3/pcodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

hit	Score	Match	Length	DB	ID	Description
1	296.6	15.9	4407	6	US-10-196-749-351	Sequence 351, App
2	284.8	15.2	12963	6	US-10-517-441-68	Sequence 68, App
3	284.4	15.2	1237	6	US-10-511-937-430	Sequence 430, App
4	284.4	15.2	2549	7	US-11-293-697-1413	Sequence 1413, App
5	282.4	15.1	3329	7	US-11-293-697-1502	Sequence 1502, App
6	271	14.5	7432	6	US-10-517-441-2	Sequence 2, App
7	270.8	14.5	2837	7	US-11-293-697-2282	Sequence 2282, App
8	270.8	14.5	394191	6	US-10-506-549-3	Sequence 3, App
9	269.6	14.4	118899	7	US-11-189-279-64	Sequence 64, App
10	269	14.4	10865	6	US-10-517-441-109	Sequence 109, App
11	266.6	14.3	7432	6	US-10-517-441-2	Sequence 2, App
12	262.2	14.0	3495	7	US-11-293-697-463	Sequence 463, App
13	260.8	13.9	2284	7	US-11-293-697-2393	Sequence 2393, App
14	257.8	13.8	3174	7	US-11-293-697-116	Sequence 116, App
15	257	13.7	2435	6	US-10-517-441-94	Sequence 94, App
16	256.8	13.7	2033	7	US-11-293-697-2350	Sequence 2350, App
17	254.4	13.6	176373	7	US-11-175-714-87	Sequence 87, App
18	253.2	13.5	11029	6	US-10-517-441-29	Sequence 29, App
19	247.4	13.2	6521	6	US-10-517-441-32	Sequence 32, App
20	242.4	13.0	152331	7	US-11-175-714-86	Sequence 86, App
21	239.8	12.8	1740	7	US-11-293-697-1833	Sequence 1833, App
22	239.2	12.8	2337	7	US-11-293-697-1071	Sequence 1071, App
23	238.8	12.8	2133	7	US-11-293-697-468	Sequence 468, App
24	237.6	12.7	2730	7	US-11-293-697-250	Sequence 250, App
25	236.4	12.6	2638	7	US-11-293-697-55	Sequence 55, App

C	26	232.6	12.4	16125	7	US-11-236-238-1	Sequence 1, App
C	27	231	12.4	2676	7	US-11-293-697-125	Sequence 125, App
C	28	230.6	12.3	2690	7	US-11-293-697-682	Sequence 682, App
C	29	228.6	12.2	3097	7	US-11-293-697-1206	Sequence 1206, App
C	30	228	12.2	56580	6	US-10-553-298-1	Sequence 1, App
C	31	225.6	12.1	2552	7	US-11-293-697-360	Sequence 360, App
C	32	222.8	11.9	1801	7	US-11-293-697-1829	Sequence 1829, App
C	33	222.8	11.9	11899	7	US-11-189-279-64	Sequence 64, App
C	34	222.4	11.9	1601	7	US-11-328-161-25	Sequence 25, App
C	35	222.4	11.9	1645	7	US-11-328-161-12	Sequence 12, App
C	36	221.4	11.8	1639	6	US-10-511-937-444	Sequence 444, App
C	37	220.6	11.8	56580	6	US-10-553-298-1	Sequence 1, App
C	38	220.4	11.8	2860	7	US-11-293-697-25	Sequence 25, App
C	39	220.2	11.8	6101	6	US-10-517-441-93	Sequence 93, App
C	40	220.2	11.8	27684	6	US-10-105-871-4	Sequence 4, App
C	41	220	11.8	2024	7	US-11-293-697-304	Sequence 304, App
C	42	217.6	11.6	2836	7	US-11-293-697-30	Sequence 30, App
C	43	217.2	11.6	11029	6	US-10-517-441-29	Sequence 29, App
C	44	217.2	11.6	128361	6	US-10-505-928-151	Sequence 151, App
C	45	216.6	11.6	2072	7	US-11-293-697-1005	Sequence 1005, App

ALIGNMENTS

RESULT 1
US-10-196-749-351
Sequence 351, Application US/10196749
Publication No. US20060094864A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Gurney, Auctin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colleen K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C340
CURRENT FILING DATE: 2002-07-16
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 351
LENGTH: 4407
TYPE: DNA
ORGANISM: Homo Sapien
US-10-196-749-351

840 TTGTATTTTATAGATAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCG 899
 1979 TTGTATTTTATAGATAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCG 2037
 900 ACCTCAGGTATCCGCTGGCTTGGCTCCCAAGTCTGGGATTCAGGCGTGAAGCCAT 959
 2038 ACCTCAGGTATTCAGCTGGCTTGGCTCCCAAGTCTGGGATTCAGGCGTGAAGCCAC 2097
 960 CGGCGCAGGCTGAGCTACTCTTAGCT 989
 2098 TGAAGCCAGGCTGCTATTGTTTAAATT 2127

1.7 8
 (-506-549-3/c
 Sequence 3, Application US/10506549
 Publication No. US20060100417A1
 GENERAL INFORMATION:
 APPLICANT: APPLERA CORPORATION
 TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
 TITLE OF INVENTION: AND USES THEREOF
 FILE REFERENCE: CLO01361-US
 CURRENT APPLICATION NUMBER: US/10/506,549
 PRIOR FILING DATE: 2004-09-03
 PRIOR APPLICATION NUMBER: 60/361,343
 PRIOR FILING DATE: 2002-03-05
 NUMBER OF SEQ ID NOS: 4
 (SOFTWARE: FASTSEQ for Windows Version 4.0
) ID NO 3
) LENGTH: 394191
) TYPE: DNA
) ORGANISM: Homo sapiens
) FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(394191)
 OTHER INFORMATION: n = A,T,C or G
 (-506-549-3

Query Match 14.5%; Score 270.8; DB 6; Length 394191;
 % Local Similarity 78.2%; Pred. No. 1.4e-34;
 % Conserved 0; Mismatches 97; Indels 8; Gaps 4;
 518 TTTCTTTTATTTTATAGATAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCG 577
 331446 TTTTATTTTATAGATAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCG 331387
 578 CTCTGCCACTGCAACTCTGCTCTCCGGAATCAAGGATTTCTGCTCAAGCTTCCCA 637
 331386 CTCTGCCACTGCAACTCTGCTCTCCGGAATCAAGGATTTCTGCTCAAGCTTCCCA 331327
 638 AGTAGGCGGATTAAGAGTGAAGGCAACCAACCCAGC-----TTTTTATTTTGGAGA 692
 331326 GGTAGTGGGACCAAGATTAAGAGTGAAGGCAACCAACCCAGC-----TTTTTATTTTGGAGA 331267
 693 CAGAGTCTTCG-CCTGTCAACCAAGGCTGG-AGTAGTGGCATTGATCTCACTGCTGCG 750
 331266 TGAAGTTTGGCTCTGTGTGGCCAGGCTGGAAGTGAAGTGTGCTTTTGTCTACCTGCA 331207
 751 ACCTGCACTTCCGGGTTCAAGCAATTTCTGCTCAAGTCTCTGATGATAGATTGA 810
 331206 ACCTGCACTTCCGGGTTCAAGCAATTTCTGCTCAAGTCTCTGATGATAGATTGA 331147
 811 CAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 869
 331146 CAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 331087
 870 CATGTTGGCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 929
 331086 CATGTTGGCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 331027
 930 CAAAGTCTGGGATTCAGGCTTGAAGCCTGAGCCAGGCTGAGCTACTCTTAGCT 989

Db 331026 CAAAGTCTGGGATTCAGGCTTGAAGCCTGAGCCAGGCTACTCTTAGCTAT 330967
 QY 990 CT 991
 Db 330966 TT 330965

RESULT 9
 US-11-189-279-64
 Sequence 64, Application US/11189279
 Publication No. US20060115829A1
 GENERAL INFORMATION:
 APPLICANT: MAO, LI
 APPLICANT: WANG, JIE
 APPLICANT: LUO, WANG
 TITLE OF INVENTION: A METHOD OF TREATING CANCER
 FILE REFERENCE: UTXC:875US
 CURRENT APPLICATION NUMBER: US/11/189,279
 PRIOR FILING DATE: 2005-07-26
 PRIOR APPLICATION NUMBER: 60/598,554
 PRIOR FILING DATE: 2004-08-03
 NUMBER OF SEQ ID NOS: 69
 (SOFTWARE: Patent Ver. 2.1
) SEQ ID NO 64
) LENGTH: 118899
) TYPE: DNA
) ORGANISM: Homo sapiens
 US-11-189-279-64

Query Match 14.4%; Score 269.6; DB 7; Length 118899;
 % Local Similarity 78.0%; Pred. No. 2.4e-34;
 % Conserved 0; Mismatches 89; Indels 19; Gaps 4;
 510 TTTCTTTTCTTTTATTTTATAGATAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCG 569
 5250 TTTTATTTTATAGATAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCG 5309
 570 GGCATGATCTCTGCGCACTGCAACTCTGCTCTCCGGAATCAAGCAATTTCTGCTGCTCA 629
 5310 GGCATGATCTCTGCGCACTGCAACTCTGCTCTCCGGAATCAAGCAATTTCTGCTGCTCA 5369
 630 GCTCCCAAGTCTGGGATTAAGAGTGAAGGCAACCAACCCAGC-----676
 5370 GCTCCCAAGTCTGGGATTAAGAGTGAAGGCAACCAACCCAGC-----676
 677 --TTTTTATTTTATAGATAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCG 5429
 5430 TTTTATTTTATAGATAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCG 5489
 734 GATCTGATCTGCTGCACTGCACTCCAGCTCCGAGTTCAAGCAATTTCTGCTGCTGCTG 793
 5490 GATCTGATCTGCTGCACTGCACTCCAGCTCCGAGTTCAAGCAATTTCTGCTGCTGCTG 5549
 794 CTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 853
 5550 CCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5607
 854 AGAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCGGAGCTGAGTATGCTG 913
 5608 AGAGATGCGCTTTTGGCATTGGCCATGCTAGCTGGAACCCCGGAGCTGAGTATGCTG 5666
 914 GCTGAGCTTGGCTCCCAAGTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 973
 5667 GCTGAGCTTGGCTCCCAAGTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5726
 974 GCTAGCTCTT 983
 Db 5727 CTCTTGTCTT 5736

RESULT 10
 US-10-517-441-109

Sequence 109, Application US/10517441
Application No. US20060121467A1

GENERAL INFORMATION:

PPLICANT: POEKENS, John
PPLICANT: HARBECK, Nadia
PPLICANT: KOENIG, Thomas
PPLICANT: MAIER, Sabine
PPLICANT: MARTENS, John
PPLICANT: MODEL, Fabian
PPLICANT: NIMWICH, Inko
PPLICANT: RUTAN, Tamas
PPLICANT: SCHMITT, Armin
PPLICANT: SCHMITT, Manfred
PPLICANT: LOOK, Maxime P.
PPLICANT: MARX, Almut
PPLICANT: HOEFER, Heinz

TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cancer

FILE OR INVENTION: proliferative disorders

FILE REFERENCE: 4676/93

CURRENT APPLICATION NUMBER: US/10/517,441

PARENT FILING DATE: 2004-12-11

FILOR APPLICATION NUMBER: PCT/EP2003/010881

FILOR FILING DATE: 2003-10-01

FILOR APPLICATION NUMBER: DE 10317955.0

FILOR FILING DATE: 2003-04-17

FILOR APPLICATION NUMBER: DE 10300096.8

FILOR FILING DATE: 2003-01-07

FILOR APPLICATION NUMBER: DE 10245779.4

FILOR FILING DATE: 2002-10-01

NUMBER OF SEQ ID NOS: 2147

ID NO 109

LENGTH: 10865

TYPE: DNA

ORGANISM: Homo Sapiens

Accession: -517-441-109

14.4%; Score 269; DB 6; Length 10865;
Local Similarity 75.4%; Pred. No. 3.8e-34; Indels 13; Gaps 3;
Conservative 0; Mismatches 110;

504 TTTTTCCTTTCTTTCTTTTCTTTTGTGAGACAGATCTGCTC-TGTCGCCCGAGCTGA 562
6664 TTTAATTAAATTATTTATTTATTTATTTTGAGTAGAGTTTCACTTATATCCAGGCTAGA 6723
563 GTGCAGTCGATGATCTCTGCCCCCATGCAACTCTGCTCCCGGATTCAGAGATTCTCC 622
6724 GTGCAGTAGGGGTATCTCGGCTCATGCAACTGCTGCTCCCGGTTCAAGAAATTTCTC 6783
623 TGCCCTGAGCTCCCAAGTAGCTGGATTTACAGGTGACGCCACACACCAGC----- 675
6784 TGCCCTGAGCTCTCGAGTAGCTGGATTTACAGGCATGTCCACACAGCTGGCTAATTTT 6843
676 TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTACCCAGGCTGAGTACAGTGCATGA 735
6844 TTTTTTTTTTTTGGAGACAGAGTCTCATCTGTGCGCCGAGGTGAGTGCATGCGCAA 6903
736 TTCACAGTTCACGTGACCTCACCTCCCGGTTCAAGCAATTCCTGTGCTGAGTCTCCT 795
6904 TCTGCGTCACTGCAAGCTTGCTGCTCCCGAGTTCACGCAATTCCTCCGACTAGCTCC 6963
796 GAGTAGCTAGGATTAAGAAAGTGCACCTCCAGCTTACGTAAT-----TTTTGTATTTT 850
6964 AAGTAGCTGGAGATCAAGAGGCGCTGCACACAATCCGGTAATTTTGATTTTGTATTTT 7023
851 AGTAGAGATCGGCTTTTGCCATGTTGGCCATGCTAATCTGGAACCCCGAGACTCAGGTGA 910
7024 AGTAGAACGGGGTTCACCGGTGTGTCAGCGTGTCTGAACTCTCAGACTCAGGTGA 7083
911 TCGCGTGGCCTTGAGCCTCCCAAAGTGTGGAGTTGAGGCGTAGAGGCATCGGCGCAGGCG 970
7084 TCTGCCCACTTCTGCTCCCAAGAGTGTGGAGTTAAGGACATGAGGCATCTGTGCCGCGC 7143
971 TAGACTACTCTTATGCTCT 991

```

RESULT 11
US-10-517-441-2
; Sequence 2, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MATER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMERICH, Inko
; APPLICANT: RUFAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almut
; APPLICANT: HOFERLER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cancer
; FILE REFERENCE: 47675-93
CURRENT APPLICATION NUMBER: US/10/517,441
CURRENT FILING DATE: 2004-12-11
PRIOR APPLICATION NUMBER: PCT/EP2003/010881
PRIOR FILING DATE: 2003-10-01
PRIOR APPLICATION NUMBER: DE 10317955.0
PRIOR FILING DATE: 2003-04-17
PRIOR APPLICATION NUMBER: DE 10300096.8
PRIOR FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: DE 10245779.4
PRIOR FILING DATE: 2002-10-01
NUMBER OF SEQ ID NOS: 2147
SEQ ID NO 2
LENGTH: 7432
TYPE: DNA
ORGANISM: Homo Sapiens
US-10-517-441-2

Query Match      14.3%; Score 266.6; DB 6; Length 7432;
Best Local Similarity 76.6%; Pred. No. 9.2e-34;
Matches 382; Conservative 0; Mismatches 104; Indels 13; Gaps 4;

QY      488   TGAGCTGCTGCTGCTTTTTTTTCTTTTCTTTTCTTTTGGAGACAGAGTCGTGCTC    547
DB      1695   TGTACGCTGTGTTTGATGCATATGCATTTTTTTTTTTTTGGAGGAGTCACTCT    1754

QY      548   GTGCCCCAGGCGTGAGTGACAGTGCAT-----GATCTGCCCCACATGCAACTCTG    598
DB      1755   GTCACCCAGGCTGAGTGACAGTGTAGCAGTGTGCGCATCTTGCTCACTGCAATCCG    1814

QY      599   CCTCCCGGATTCAAGCGATTCTCCGTGCTCAGCCTCCCAAAGTAGCTGGATTACAGGTGC    658
DB      1815   CCTACCAGTTCAAGGAGATTCTCTGCTCAGTCTTCAGAGTATTTGGGACATAAGACAC    1874

QY      659   ACGCCACACACCCAGCT-TTTTTTATTTGGAGACAGAGTCCTGCCCTGATGACCCAGGC    717
DB      1875   ACGCCACACATGCTGGGCTATTTTTTTTTTTTGAAGACGAGTCTGCTCTGTATCCAGGC    1934

QY      718   TGAAGTACAGTGGCATATCTCAGTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATT    777
DB      1935   TGAAGTACAGTGGGCTATCTTGCTCACTGCAACGATCCGCTCCGGGGTTCAAGCATTT    1994

QY      778   CTCTGCTCAGTCTCTGAGTAGCTAGATTACAGAAGTGCACCTCCAGCTTCACTTA    837
DB      1995   CTCTGCTCAGCTCTCCCGAGTAGCTGGGACATAAGAGGCCAACACACGCTGAGCTAA    2054

QY      838   -TTTTTGAATTTTAGTAGAGATGCGCTTTTGGCATGTGGCCATGCTAGCTGAACCC    896
DB      2055   TTTTGTGATTTTACTAGAGACGGGGTTTCCGCTGTGATTTACCAAGAGATGATCTCATATC    2114

```

897 CGACCTCAGAGGATGATCCGCTGGCCCTGGCCCTCCCAAAGTCTGGAGTTTCAGAGCCTGAGC 956
2115 CTGACCTC-CTTGATCCGCTGGCCCTCGGCTTCCAAAGTCTAGATTTATAGGCGCTGAGC 2172
957 CATCGCGCCAGGCGCTGAGC 975
2173 CACTCGCGCTGGCCCAATGC 2191

IT 12
 -293-697-463
 Sequence 463, Application US/11293697
 Publication NO. US20060105376A1
 SERIAL INFORMATION:
 APPLICANT: HELIX RESEARCH INSTITUTE
 TITLE OF INVENTION: NOVEL full length CDNA
 SE REFERENCES: H1-A0106
 CURRENT APPLICATION NUMBER: US/11/293,697
 CURRENT FILING DATE: 2005-12-05
 PRIOR APPLICATION NUMBER: US/10/108,260
 PRIOR FILING DATE: 2002-03-28
 NUMBER OF SEQ ID NOS: 5458
 SOFTWARE: PatentIn Ver. 2.1
 1) ID NO. 463
 LENGTH: 3495
 TYPE: DNA
 ORGANISM: Homo sapiens
 -293-697-463

Query Match 14.0%; Score 262.2; DB 7; Length 3495;
 Percent Local Similarity 75.6%; Pred. No. 4.7e-33;
 Mismatches 377; Conservative 0; Mismatches 93; Indels 29; Gaps 3;

[illegible]

1,7 13
..-293-697-2393

```

Sequence 2393, Application US/11293697
Publication No. US20060105376A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/11/293,697
CURRENT FILING DATE: 2005-12-05
PRIOR APPLICATION NUMBER: US/10/108,260
PRIOR FILING DATE: 2002-03-28
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2393
LENGTH: 2284
TYPE: DNA
ORGANISM: Homo sapiens
US-11-293-697-2393

Query Match      13.9%   Score 260.8; DB 7; Length 2284;
Best Local Similarity 77.3%;   Pred. No. 8,1e-33;
Matches 377; Conservative 0; Mismatches 87; Indels 24; Gaps 4;
```

[illegible]

RESULT 14
 US-11-293-697-116
 Sequence 116, Application US/11293697
 Publication No. US20060105376A1
 GENERAL INFORMATION
 APPLICANT: HELIX RESEARCH INSTITUTE
 TITLE OF INVENTION: Novel full length cDNA
 FILE REFERENCE: H1-A0106
 CURRENT APPLICATION NUMBER: US/11/293,697
 CURRENT FILING DATE: 2005-12-05
 PRIOR APPLICATION NUMBER: US/10/108,260
 PRIOR FILING DATE: 2002-03-28
 NUMBER OF SEQ ID NOS: 5458
 SOFTWARE: Patent Ver. 2.1

